

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 11, 2005, 10:08:01 ; Search time 217 Seconds
(without alignments)
3804.014 Million cell updates/sec

Title: US-10-783-417-1
Perfect score: 2208
Sequence: 1 atgatacaataacagataa.....atcatacaaaacacttga 2208

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 3392430 seqs, 186927314 residues

Total number of hits satisfying chosen parameters: 6784860

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

Published Applications NA New:
1: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq.*
2: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq.*
3: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq.*
4: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq.*
5: /cgn2_6/ptodata/2/pubpna/FCI_NEW_PUB.seq.*
6: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq.*
7: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq.*
8: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq2.*
9: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq3.*
10: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq.*

Prod. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	69	3.1	173602	7	US-11-121-086-25
2	63.8	2.9	171486	7	US-11-121-086-105
3	60.8	2.8	2031	7	US-11-058-727-51
4	60.8	2.8	2031	7	US-11-058-727-83
5	60.8	2.8	2031	7	US-11-108-389-51
6	60.8	2.8	2031	7	US-11-108-389-83
7	60.6	2.7	173602	7	US-11-121-086-25
8	59.6	2.7	2028	7	US-11-058-727-39
9	59.6	2.7	2028	7	US-11-058-727-71
10	59.6	2.7	2028	7	US-11-108-389-39
11	59.6	2.7	2028	7	US-11-108-389-71
12	58.2	2.6	2022	7	US-11-058-727-13
13	58.2	2.6	2022	7	US-11-108-389-13
14	58.2	2.6	3633	7	US-11-058-727-3
15	58.2	2.6	3633	7	US-11-108-389-3
16	58.2	2.6	6613	7	US-11-058-727-18
17	58.2	2.6	6613	7	US-11-108-389-18
18	58.2	2.6	139054	7	US-11-121-086-96
19	58	2.6	2025	7	US-11-058-727-41
20	58	2.6	2025	7	US-11-058-727-73
21	58	2.6	2025	7	US-11-108-389-41
22	58	2.6	2025	7	US-11-108-389-73
23	57.6	2.6	2019	7	US-11-058-727-55

24	57.6	2.6	2019	7	US-11-058-727-59	Sequence 59, Appl
25	57.6	2.6	2019	7	US-11-058-727-87	Sequence 87, Appl
26	57.6	2.6	2019	7	US-11-058-727-91	Sequence 91, Appl
27	57.6	2.6	2019	7	US-11-108-389-55	Sequence 55, Appl
28	57.6	2.6	2019	7	US-11-108-389-59	Sequence 59, Appl
29	57.6	2.6	2019	7	US-11-108-389-87	Sequence 87, Appl
30	57.6	2.6	2019	7	US-11-108-389-91	Sequence 91, Appl
31	57	2.6	171486	7	US-11-121-086-105	Sequence 105, Appl
32	56.8	2.6	151169	7	US-11-121-086-38	Sequence 38, Appl
33	55.2	2.5	2022	7	US-11-058-727-49	Sequence 49, Appl
34	55.2	2.5	2022	7	US-11-058-727-81	Sequence 81, Appl
35	55.2	2.5	2022	7	US-11-108-389-49	Sequence 49, Appl
36	55.2	2.5	2022	7	US-11-108-389-81	Sequence 81, Appl
37	54.8	2.5	2025	7	US-11-058-727-47	Sequence 47, Appl
38	54.8	2.5	2025	7	US-11-058-727-79	Sequence 79, Appl
39	54.8	2.5	2025	7	US-11-108-389-47	Sequence 47, Appl
40	54.8	2.5	2025	7	US-11-108-389-79	Sequence 79, Appl
41	54.4	2.5	2019	7	US-11-058-727-57	Sequence 57, Appl
42	54.4	2.5	2019	7	US-11-058-727-89	Sequence 89, Appl
43	54.4	2.5	2019	7	US-11-108-389-57	Sequence 57, Appl
44	54.4	2.5	2019	7	US-11-108-389-89	Sequence 89, Appl
45	54.4	2.5	2022	7	US-11-058-727-21	Sequence 21, Appl

ALIGNMENTS

RESULT 1									
US-11-121-086-25/c									
; Sequence 25, Application US/11121086									
; Publication No. US20050266459A1									
; GENERAL INFORMATION:									
; APPLICANT: NIELSEN, KIRSTEN V.									
; TITLE OF INVENTION: NUCLEIC ACID PROBES AND NUCLEIC ACID ANALOG PROBES									
; FILE REFERENCE: 09138.6000-00000									
; CURRENT FILING DATE: 2005-05-04									
; PRIOR APPLICATION NUMBER: 60/567,570									
; PRIOR FILING DATE: 2004-05-04									
; NUMBER OF SEQ ID NOS: 107									
; SOFTWARE: PatentIn version 3.3									
; SEQ ID NO 25									
; LENGTH: 173602									
; TYPE: DNA									
; ORGANISM: Homo sapiens									
US-11-121-086-25									
Query Match 3.1%; Score 69; DB 7; Length 173602;									
Best Local Similarity 45.7%; Pred. No. 0.0013;									
Matches 430; Conservative 0; Mismatches 500; Indels 11; Gaps 5;									
QY	585	TTAAATACGATTGGAGATGTCACATGATTTATTCGAGAAATACCTGTTCCAACT	644						
DB	136987	TAT	136928						
QY	645	TGAACCTATATAACGCTATACCTATTTATGCGGAGCTGATTTTCATTTAA	704						
DB	136927	AATATATTAACATTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA	136668						
QY	705	TTTATTAACAAGGTCGATGCTGATGATGAATGGAATGCAGATATACATCTTCA	764						
DB	136867	TAT	136608						
QY	765	AATGAACCTATATGCTGGAACATCAGATGACTATTTAAACTTTAAAGAAATATACC	824						
DB	136807	AATATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA	136748						
QY	825	TAAATATATGTAACATTTGTCGAATTAATTAATTAATTAATTAATTAATTAATTAATTA	884						
DB	136747	TAT	136688						
QY	885	ACCAATATGAATGAGTATATTTAATGACTATGAGATATATGACCATTAATCTGATTT	944						

[illegible]

```

RESULT 2
US-11-121-086-105
; Sequence 105, Application US/11121086
; Publication No. US20050266459A1
;
GENERAL INFORMATION:
;
APPLICANT: POULSEN, TIM S.
;
APPLICANT: NIELSEN, KIRSTEN V.
;
TITLE OF INVENTION: NUCLEIC ACID PROBES AND NUCLEIC ACID ANALOG PROBES
;
FILE REFERENCE: 09138.6000-00000
;
CURRENT APPLICATION NUMBER: US/11/121.086
;
CURRENT FILING DATE: 2005-05-04
;
PRIOR APPLICATION NUMBER: 60/567,570
;
PRIOR FILING DATE: 2004-05-04
;
NUMBER OF SEQ ID NOS: 107
;
SOFTWARE: PatentIn version 3.3
;
SEQ ID NO 105
;
LENGTH: 171486
;
TYPE: DNA
;
ORGANISM: Homo sapiens
;
US-11-121-086-105

```

Query Match	2.9%;	Score 63.8;	DB 7;	Length 171486;
Best Local Similarity	44.0%;	Pred. No. 0.011;		
Matches 410;	Conservative 0;	Mismatches 512;	Indels 9;	

QY 690 TAATTTCATTAAATTTATTTACAACAAGTGTGAATTCGTGATGAATGSAATCCAGA 749
 |||||
 Db 16132 TAAATATATTTATTTATATAATATATATCTTATATATTTATATATATATATATATATATAT 161919
 |||||

Oy		750	TATACATCCCTGCAAAATTGAACCTTAATGCGGAAACATGCAGATGACTATTAATAAACCCTTT	809
Db		16192	ATTATATGTAT	16251
Oy		810	AAAAAGAAAATATACCTTAAATATATAGTACTATTGTGCAAAATACCTATAGAACAAGACTAAA	869
Db		16252	TAT	16311
Oy		870	AATATCTTAGAGCAGAACCAATATGGAATGGAGTATAT-TTAAATGACTATCCGAAGTATAT	928
Db		- 16312	AACAT	16371
Oy		929	TGACCATTAICTGTATATAGATATACCACATCTCAATTTTCCTTATATGATATATAAAAGATATA	988
Db		16372	TTAT	16431
Oy		989	GAGATTCATATGAGCAGATATGAAAGTAAAAAGCCTTAAAGATGACCTCACAAAGAAATTT	1048
Db		16432	TAT	16491
Oy		1049	ATACAACTGMAATPAAATTTTGATCGATCTCCATCACTTAGAGTCAACCCAATCTAGCTA	1108
Db		16492	ATTATCATATTAAT	16551
Oy		1109	CGATGATATATATATTAACACGTGCAAGTTTTAAATTAATTTTCATTTTGAACAATTTTA	1168
Db		16552	AT	16611
Oy		1169	TTTTTTATATACAGAA-ATACAAATTTCCGGAAATCGTTTGTTGGTATTTCTATATGCTGAT	1227
Db		16612	TAAAT	16671
Oy		1228	GCACTACTATATAGCAATACTATATACATGAACTTTATATAGGAAAAGAACAGTTTCACCCC	1287
Db		16672	TAT	16724
Oy		1288	ACAAACAAAACATTAAGACACTTTGAATCTTAAAGTTTCAAATTTCTATCTGATAGACAA	1347
Db		16725	ATTAAT	16784
Oy		1348	TCACTCTCTGTTTTCCCCATATTCACACACACTTATATATATATATATATATATATATATATTA	1407
Db		16785	TTTACAT	16844
Oy		1408	AATGGCTCATCTAACACACACCTCAAATTTTCAGCAGGAGGGCTTTATCTATATATCA	1467
Db		16845	TAT	16904
Oy		1468	AACACAACTTTTTTTCAAATTTCTTAGAAAATAAAGCTGCATATCTAGTATATATCCAGGT	1527
Db		16905	TAAATATATATTTGATATATTTATATATGATATGATATATATATATATATATATATATATATAT	16964
Oy		1528	TGTTCCAAAACCTTATATATACATATAGTCATATTTATCCCAATTTTTCAATTAATTTACTAT	1587
Db		16965	ATGAT	17024
Oy		1588	TCTTATATGATATGGATTTACGCTACAAATAT	1618
Db		17025	TTATATATATATATATATATATATATATATAT	17055

RESULT 3
US-11-058-727-51
; Sequence 51, Application US/11058727
; Publication No. US20050261483A1

/	APPLICANT:	Ronald D. Flannagan
/	APPLICANT:	Rafael Herrmann
/	APPLICANT:	Theodore W. Kahn
/	APPLICANT:	Albert L. Lu
/	APPLICANT:	Billy Fred McCutchen
/	APPLICANT:	James K. Presnail

```

; APPLICANT: James F.H. Wong
; APPLICANT: Cao-Guo Yu
; TITLE OF INVENTION: Genes Encoding Proteins with Pesticidal
; TITLE OF INVENTION: Activity
; FILE REFERENCE: 35718/287809
; CURRENT APPLICATION NUMBER: US/11/058,727
; CURRENT FILING DATE: 2005-02-15
; PRIOR APPLICATION NUMBER: 60/391,786
; PRIOR FILING DATE: 2002-06-26
; PRIOR APPLICATION NUMBER: 60/460,787
; PRIOR FILING DATE: 2003-04-04
; PRIOR APPLICATION NUMBER: 10/606,320
; PRIOR FILING DATE: 2003-06-25
; NUMBER OF SEQ ID NOS: 134
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO: 51
; LENGTH: 2031
; TYPE: DNA
; ORGANISM: Bacillus thuringiensis (mutated)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(2031)
; US-11-058-727-51

```

```

Query Match      2.8%; Score 60.8; DB 7; Length 2031;
Best Local Similarity 47.1%; Pred. No. 0.014;
Matches 292; Conservative 0; Mismatches 292; Indels 36; Gaps 2;

```

```

QY 367 GACAAAACGATGACACAACTTTTAAATGGAGAAATTTTGTGATACACCGTTA 426
DB 332 GAAAAGATGCAATGGAATTTTATGAAACAAGAGAAAGAACTCAATTAATCAAAAATA 381
QY 427 ACAGAAAGATGAAACAGCTAAAGTTACAACTTTAGAGAGATTTAGCAAAATATTACA 486
DB 382 GCAGAAATGCAAGAAATTAAGCGCTTTGGAATTAAGAGATTAAGTAATTAATTA 441
QY 487 AGCTAATATACAGCATTTAGATGATTGAGAAATTTAAAGAACTCAAGCTCTGATTA 546
DB 442 TTATATCTAATCGCGCTTGAAGAAATGGAAGAAATTCATTTGCAAGTGAAGTATTT 501
QY 547 CCACCATCATGACATTACACAGAGCTGCTGACTCTTAAATACGATTGAGAAATGTT 606
DB 502 AATGCTCCGCGCCAGCCTTACGAGATG-----TGCAGAAATCGATTGAAATCCCTG 552
QY 607 CACATGATTTTATTCGAGAAATACCTGTTTCAACTGGAATTAATTAAGCTATTA 666
DB 553 GATAGTTTATTTACCAATTAATGCGCATCTTTTAAAGTGAACAATTTGAACTACCATTC 612
QY 667 CTACCTATTTATGCGCAAGCTGCTAATTTTCATTTAAATTTTATTAACAAGTGTGAA 726
DB 613 CTTACTGATATGCAATGCGCAACCTTCATTTTACGTATTTAAAGAGCGGTCAATT 672
QY 727 TTGCTGATGAAATGCAATGCAATATACATCTTCAAAATTTGAACCTTAATGCTGGA 786
DB 673 TTTGAGAGAAATGAGGATGTCACAACT-----ACT 705
QY 787 TCAGATGCTATTTAATCTTTTAAAGAAATATTAATTAATTAATTAATTAATTAATTA 846
DB 706 ATTAATTAATTAATTAATGATGTCAAATGAAATTAATTAATTAATTAATTAATTAATTA 765
QY 847 AATACCTATAGAACAGAACTAATAAATCTTAGAGAGAAACCAATTAATTAATTAATTA 906
DB 766 AAGTGATGAAATGAGTGTATTAACAATAAAGCAAGCGCTTAACAATGAGGTGAC 825
QY 907 TTTAATGATATGAGATATATGACATTAATTAATTAATTAATTAATTAATTAATTAATTT 966
DB 826 TATAACCAATTCGTAAGAGAAATGACATGCGGTTTATGATGTTGTCATTAATTTCCA 885
QY 967 TTATATGATATTAATAAGATA 986
DB 886 AATTATGACACAGCAGTA 905

```

```

RESULT 4
US-11-058-727-83
; Sequence 83, Application US/11058727
; Publication No. US20050261483A1
; GENERAL INFORMATION:
; APPLICANT: Andre R. Abad
; APPLICANT: Ronald D. Flanagan
; APPLICANT: Rafael Herzmann
; APPLICANT: Theodore W. Kahn
; APPLICANT: Albert L. Lu
; APPLICANT: Billy Fred McCutchen
; APPLICANT: James K. Preenail
; APPLICANT: James F.H. Wong
; APPLICANT: Cao-Guo Yu
; TITLE OF INVENTION: Genes Encoding Proteins with Pesticidal
; TITLE OF INVENTION: Activity
; FILE REFERENCE: 35718/287809
; CURRENT APPLICATION NUMBER: US/11/058,727
; CURRENT FILING DATE: 2005-02-15
; PRIOR APPLICATION NUMBER: 60/391,786
; PRIOR FILING DATE: 2002-06-26
; PRIOR APPLICATION NUMBER: 60/460,787
; PRIOR FILING DATE: 2003-04-04
; PRIOR APPLICATION NUMBER: 10/606,320
; PRIOR FILING DATE: 2003-06-25
; NUMBER OF SEQ ID NOS: 134
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO: 83
; LENGTH: 2031
; TYPE: DNA
; ORGANISM: Bacillus thuringiensis
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(2031)
; US-11-058-727-83

```

```

Query Match      2.8%; Score 60.8; DB 7; Length 2031;
Best Local Similarity 47.1%; Pred. No. 0.014;
Matches 292; Conservative 0; Mismatches 292; Indels 36; Gaps 2;

```

```

QY 367 GACAAAACGATGACACAACTTTTAAATGGAGAAATTTTGTGATACACCGTTA 426
DB 332 GAAAAGATGCAATGGAATTTTATGAAACAAGAGAAAGAACTCAATTAATCAAAAATA 381
QY 427 ACAGAAACATTAACAGCTAAAGTTACAACTTTAGAGAGATTTAGCAAAATATTACA 486
DB 382 GCAGAAATGCAAGAAATTAAGCGCTTTGGAATTAAGAGATTAAGTAATTAATTA 441
QY 487 AGCTAATATACAGCATTTAGATGATTGAGAAATTTAAAGAACTCAAGCTCTGATTA 546
DB 442 TTATATCTAATCGCGCTTGAAGAAATGGAAGAAATTCATTTGCAAGTGAAGTATTT 501
QY 547 CCACCATCATGACATTACACAGAGCTGCTGACTCTTAAATACGATTGAGAAATGTT 606
DB 502 AATGCTCCGCGCCAGCCTTACGAGATG-----TGCAGAAATCGATTGAAATCCCTG 552
QY 607 CACATGATTTTATTCGAGAAATACCTGTTTCAACTGGAATTAATTAATTAATTAATTA 666
DB 553 GATAGTTTATTTACCAATTAATGCGCATCTTTTAAAGTGAACAATTTGAACTACCATTC 612
QY 667 CTACCTATTTATGCGCAAGCTGCTAATTTTCATTTAAATTTTATTAACAAGTGTGAA 726
DB 613 CTTACTGATATGCAATGCGCAACCTTCATTTTACGTATTTAAAGAGCGGTCAATT 672
QY 727 TTGCTGATGAAATGCAATGCAATATACATCTTCAAAATTTGAACCTTAATGCTGGA 786
DB 673 TTTGAGAGAAATGAGGATGTCACAACT-----ACT 705
QY 787 TCAGATGCTATTTAATCTTTTAAAGAAATATTAATTAATTAATTAATTAATTAATTA 846
DB 706 ATTAATTAATTAATTAATGATGTCAAATGAAATTAATTAATTAATTAATTAATTAATTA 765
QY 847 AATACCTATAGAACAGAACTAATAAATCTTAGAGAGAAACCAATTAATTAATTAATTA 906

```

Db 766 AAGTGATGAACTGTTTGGCAAAATTAAGGACAGCGCTTAAACAAATGGTTTAC 825
Qy 907 TTTAATGACTATCGAAGATATATGACCAATTAATGATAGATACCATCTCAATTTCT 966
Db 826 TATAACCAATTCGTAAGAAATGACACTGGCGGTTTAAATGTTTGATTAATTTCCCA 885
Qy 967 TTAATGATATTAATAAGATA 986
Db 886 AATTATGACACACGACGTA 905

RESULT 5

US-11-108-389-51
; Sequence 51, Application US/11108389
; Publication No. US20050261188A1
; GENERAL INFORMATION:
; APPLICANT: Andre R. Abad
; APPLICANT: Ronald D. Flannagan
; APPLICANT: Rafael Herrmann
; APPLICANT: Theodore W. Kahn
; APPLICANT: Albert L. Lu
; APPLICANT: Billy Fred McCutchen
; APPLICANT: James K. Presnall
; APPLICANT: James F. H. Wong
; APPLICANT: Cao-Guo Yu
; TITLE OF INVENTION: Genes Encoding Proteins With Pesticidal
; TITLE OF INVENTION: Activity
; FILE REFERENCE: 35718/291049
; CURRENT APPLICATION NUMBER: US/11/108,389
; CURRENT FILING DATE: 2005-04-18
; PRIOR APPLICATION NUMBER: 60/391,786
; PRIOR FILING DATE: 2002-06-26
; PRIOR APPLICATION NUMBER: 60/460,787
; PRIOR FILING DATE: 2003-04-04
; PRIOR APPLICATION NUMBER: 10/606,320
; PRIOR FILING DATE: 2003-06-25
; NUMBER OF SEQ ID NOS: 134
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 51
; LENGTH: 2031
; TYPE: DNA
; ORGANISM: Bacillus thuringiensis (mutated)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(2031)
US-11-108-389-51

Query Match 2.8%; Score 60.8; DB 7; Length 2031;

Best Local Similarity 47.1%; Pred. No. 0.014; Mismatches 292; Indels 36; Gaps 2;

Db 367 GACAAAGCATGTGACACAAATTTTAAATGGAGAAATTTTGTGTATACCGTTA 426
Qy 322 GAAAAGATGATGGAAATTTTATGAAACAAGTAAGAACTATTAATCAAAAAATA 381
Db 427 ACAGAAAGCATTAACACGCTAAAGTTACAACTTTAGAGATTTAGCAAAATATTACA 486
Qy 382 GCAGAAATATGCAAGAAATTAAGCGCTTCGAAATTAGAAGTTAGTAATTAACAA 441
Db 487 AGCTAATATACGACTATGATGATTTGAGAAAAATTAAGAACTCAAGCTCTCGATTA 546
Qy 442 TTAATCTTAACCTGCGCTTGAAGAAATGCAATTTGCAAGTCGAGTATTTA 501
Db 547 CCACCATATATGACATTAACAAGCTGCTGACTTTAAATATGATTTGAGAAATGT 606
Qy 502 AATGATTTCCGCGCAGCTTACGAAATG-----TGCAGAAATGCAATTTGAAATCCG 552
Db 607 CACAATGATTTTATTCGAAATATACCTGTTTCAACTGGAATTTAATAACGCTATTA 666
Qy 553 GATAGTTTATTTAAGCAATATATGCAATCTTTTAAGAGTGAACAAATTTGAAGTACATTC 612
Db 667 CTACCTAATTTATGCGCAAGCTGCTAATTTTCAATTTAATTTAACAACAGGTGCTGA 726

Db 613 CTACTGTATATATGCAATATGACAGCCACCTTCATTTACTGTTATTAAGAGACGCTCAATT 672
Qy 727 TTGGCTGATGAATGGAATGCAATATATATATATATATATATATATATATATATATAT 786
Db 673 TTGGAGAAAGATGGGATGCTCAACACT-----ACT 705
Qy 787 TCAGATGACTATTTATTAATCTTTTAAAGAAATATATATATATATATATATATATATAT 846
Db 706 AATTAAT 765
Qy 847 AATACCTATATGAACAGCACTTAAATCTTAAAGAGAGACCAATATATATATATATATAT 906
Db 766 AAGTGATGAACTGTTTGGCAAAATTAAGGACAGCGCTTAAACAAATGGTTTAC 825
Qy 907 TTTAATGACTATCGAAGATATATGACCAATTAATGATAGATACCATCTCAATTTTCT 966
Db 826 TATAACCAATTCGTAAGAAATGACACTGGCGGTTTAAATGTTTGATTAATTTCCCA 885
Qy 967 TTAATGATATTAATAAGATA 986
Db 886 AATTATGACACACGACGTA 905

RESULT 6

US-11-108-389-83
; Sequence 83, Application US/11108389
; Publication No. US20050261188A1
; GENERAL INFORMATION:
; APPLICANT: Andre R. Abad
; APPLICANT: Ronald D. Flannagan
; APPLICANT: Rafael Herrmann
; APPLICANT: Theodore W. Kahn
; APPLICANT: Albert L. Lu
; APPLICANT: Billy Fred McCutchen
; APPLICANT: James K. Presnall
; APPLICANT: James F. H. Wong
; APPLICANT: Cao-Guo Yu
; TITLE OF INVENTION: Genes Encoding Proteins With Pesticidal
; TITLE OF INVENTION: Activity
; FILE REFERENCE: 35718/291049
; CURRENT APPLICATION NUMBER: US/11/108,389
; CURRENT FILING DATE: 2005-04-18
; PRIOR APPLICATION NUMBER: 60/391,786
; PRIOR FILING DATE: 2002-06-26
; PRIOR APPLICATION NUMBER: 60/460,787
; PRIOR FILING DATE: 2003-04-04
; PRIOR APPLICATION NUMBER: 10/606,320
; PRIOR FILING DATE: 2003-06-25
; NUMBER OF SEQ ID NOS: 134
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 83
; LENGTH: 2031
; TYPE: DNA
; ORGANISM: Bacillus thuringiensis
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(2031)
US-11-108-389-83

Query Match 2.8%; Score 60.8; DB 7; Length 2031;

Best Local Similarity 47.1%; Pred. No. 0.014; Mismatches 292; Indels 36; Gaps 2;

Db 367 GACAAAGCATGTGACACAAATTTTAAATGGAGAAATTTTGTGTATACCGTTA 426
Qy 322 GAAAAGATGATGGAAATTTTATGAAACAAGTAAGAACTATTAATCAAAAAATA 381
Db 427 ACAGAAAGCATTAACACGCTTAAAGTTACAACTTTAGAGATTTAGCAAAATATTACA 486
Qy 382 GCAGAAATATGCAAGAAATTAAGCGCTTCGAAATTAGAAGTTAGTAATTAACAA 441
Db 487 AGCTAATATACGACTATGATGATTTGAGAAAAATTAAGAACTCAAGCTCTCGATTA 546

Db	442	TTATATCTA	CTACGCTT	GAGAAAGT	GGAGAAAT	CCATTTCC	AGAGCTG	CGAGTA	AGTTTA	501
Qy	547	CCACCA	TCATCAG	CAATTA	CAACAAG	CTGCCTT	GA	CTTTAAAT	ATACGATTTG	606
Db	502	AATGTTCC	CGGCA	CGCTTAC	GAGATG	-----	TG	CGAAAT	CGATTTGA	552
Qy	607	CACAATG	ATTTTAT	TCGAAAT	TACTGGTTT	CCAACTT	GAAAC	CTTATATA	AAAGCTATTA	666
Db	553	GATAGTTT	ATTTACG	AAATAT	ATGCAATCTT	TTT	TAGAGT	ACAAATTTT	TTGGAATTC	612
Qy	667	CTACCTAT	TTTATGCG	CAAGCTG	CTCTAATTT	CAATTAA	ATTATT	ATTACA	CAAGTGC	726
Db	613	CTTACTG	TATATG	CAATG	GCACCA	CTTCACTT	ACTGTT	ATTATA	AGAGCGCT	672
Qy	727	TTGCTG	ATGAATG	GAATG	CAATAT	TACATCTT	CAACAAT	TGAACCT	TATGCTG	786
Db	673	TTTGGAGA	AGAAATG	GGGATG	CTCAACA	CT-----	ACT	-----	-----	705
Qy	787	TCAGATG	ACATAT	TATTA	ACTTTT	TAAAGAA	AATAT	ACTTAAT	TATAGTAT	846
Db	706	ATTAAAT	ACATAT	TATGATG	CTCAAT	TGAACCTT	ACTG	CAATAT	TCTGATCA	765
Qy	847	AATACCT	ATAGA	CAGCACT	TAATAAA	CTTTAG	AGACGA	ACCAAT	TATGAAT	906
Db	766	AATG	GTATGA	AACGTG	TTAG	CAAAAT	TAAAG	CAAGAC	GCCTTAA	825
Qy	907	TTTATG	ACTATG	AAATAT	TGACCA	TTACTG	TATAG	TATAC	ATCTCA	966
Db	826	TATTA	CCAAAT	TCGGT	GAGAAAT	TGACACTG	GGGTTT	AGATG	TGTGCTA	885
Qy	967	TTTATG	ATATA	AAAGATA	-----	-----	-----	-----	-----	986
Db	886	AATTATG	ACACG	CACGTA	905	-----	-----	-----	-----	905

```

RESULT 7
US-11-121-086-25
; Sequence 25, Application US/11121,086
; Publication No. US2005026459A1
; GENERAL INFORMATION:
; APPLICANT: POULSEN, TIM S.
; APPLICANT: NIELSEN, KIRSTEN V.
; TITLE OF INVENTION: NUCLEIC ACID PROBES AND NUCLEIC ACID ANALOG PROBES
; FILE REFERENCE: 09138.6000-00000
; CURRENT APPLICATION NUMBER: US/11,121,086
; CURRENT FILING DATE: 2005-05-04
; PRIOR APPLICATION NUMBER: 60/567,570
; PRIOR FILING DATE: 2004-05-04
; NUMBER OF SEQ ID NOS: 107
; SOFTWARE: Patentin version 3.3
; SEQ ID NO 25
; LENGTH: 173602
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-11-121-086-25

```

Query Match	2.7%	Score 60.6;	DB 7;	Length 173602;
Best Local Similarity	43.3%	Pred. No. 0.04;		
Matches 487; Conservative	0;	Mismatches 629;	Indels 9;	Gaps 4

[illegible][illegible]

RESULT 8
US-11-056-727-39
Sequence 39, Application US/11056722-39
Publication No. US20050261483A1
GENERAL INFORMATION:
APPLICANT: Andre R. Abad
APPLICANT: Ronald D. Pinnagan
APPLICANT: Rafael Herrmann

APPLICANT: Theodore W. Kahn
APPLICANT: Albert L. Lu
APPLICANT: Billy Fred McCutchen
APPLICANT: James K. Presnail
APPLICANT: James F.H. Wong
APPLICANT: Cao-Guo Yu
TITLE OF INVENTION: Genes Encoding Proteins With Peestcidai
TITLE OF INVENTION: Activity
FILE REFERENCE: 35718/287809
CURRENT APPLICATION NUMBER: US/11/058,727
CURRENT FILING DATE: 2005-02-15
PRIOR APPLICATION NUMBER: 60/391,786
PRIOR FILING DATE: 2002-06-26
PRIOR APPLICATION NUMBER: 60/460,787
PRIOR FILING DATE: 2003-04-04
PRIOR APPLICATION NUMBER: 10/606,320
PRIOR FILING DATE: 2003-06-25
NUMBER OF SEQ ID NOS: 134
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 39
LENGTH: 2028
TYPE: DNA
ORGANISM: Bacillus thuringiensis (mutated)
FEATURES:
NAME/KEY: CDS
LOCATION: (1)...(2028)
US-11-058-727-39

Query Match 2.7%; Score 59.6; DB 7; Length 2028;
Best Local Similarity 47.1%; Pred. No. 0.023;

Matches 292; Conservative 0; Mismatches 289; Indels 39; Gaps 2;

367 GACAAAACAGTATGACACAAATTTTAAATGGAGAAATTTTGTGTATACCGTTA 426
322 GAAAAGTCAATGGAGAAATTTTATGAAACAAGTAAAGAACTCATTAATCAAAAAATA 381
427 ACAGAAAGCATTAACAGCTAAAGTTACAACTTTAGAGAGTTTACCAATTTACAA 486
382 GCAGAAATATGCAAGAAATTAAGCGCTTTCGAAATTTAGAAAGATTAGTAAATTAACCA 441
487 AGCTAATATACAGCATTAATGATGATTTGAGAAATTTAAAGATCAAGCTCCGTGATTA 546
442 TTATATCTAATCGCTGCTTGAAGATGGAAGAAATTCATTTCCAGTCGAGGTTTTCGA 501
547 CCACCATCATGACATTAACAAGCTGCTTGAATCTTAAATTAACAATTTGAGAAATGTT 606
502 AGTGAAGGTCAAGCTTACGAGA-----TGTGCAAAATCGATTTGAATCCTG 549
607 CACAAATGTTTATTTGCAAGAAATACCTGTTCCAACTTGAATCTTAATAACGCTATTA 666
550 GATAGTTTATTTACGCAATATATGCAATCTTTAGAGTGAACAAATTTGAAGTACCAATTC 609
667 CTACCTATTTATGCGCAAGCTGCTAATTTTCAATTTAATTTATTAACAAGAGTCTGAA 726
610 CTACTGTTATATGCAATGCGCAAGCTTCACTTATTTACTTTATTAAGAGCGCTCAATT 669
727 TTGCTGATGATGAGATGCAATATACATCTTCAAAATTTGAACCTTAATGCTGGAACA 786
670 TTGGAAGAGATGGGATGTCACAACT-----ACT 702
787 TCAGATGATCTATTAATCTTTAAAGAAATATACCTAAATATAGTAATCTATTTGCA 846
703 ATTAATACTATTAATGATGTCATTAATGAACTTACTGAGAAATTTCTGATCAGCTGTA 762
847 AATACCTATTAAGACGACATAAAAACTTGAAGCAACCAATATGAATGAGATATA 906
763 AAGTGTATGAAATCGTTTACCAAAATTTAAAGCAGAGCGCTTAAACAATGGCTGAC 822
907 TTTATATGATATGAGATATATATGCAATTAATGATTAATGATACATCTCAATTTTCT 966
823 TATATCAAAATTCGATGAGAAATGACACTGGCGGTTTAAAGATGTTGTTGCAATTTCCCA 882
967 TTAATATGATATTAATAAGATA 986

Db 883 AATATGACACAGCAGCTA 902

RESULT 9
US-11-058-727-71

Sequence 71, Application US/11058727
Publication No. US20050261483A1

GENERAL INFORMATION:

APPLICANT: Andre R. Abad

APPLICANT: Ronald D. Flannagan

APPLICANT: Rafael Herrmann

APPLICANT: Theodore W. Kahn

APPLICANT: Albert L. Lu

APPLICANT: Billy Fred McCutchen

APPLICANT: James K. Presnail

APPLICANT: James F.H. Wong

APPLICANT: Cao-Guo Yu

TITLE OF INVENTION: Genes Encoding Proteins With Peestcidai

TITLE OF INVENTION: Activity

FILE REFERENCE: 35718/287809

CURRENT APPLICATION NUMBER: US/11/058,727

CURRENT FILING DATE: 2005-02-15

PRIOR APPLICATION NUMBER: 60/391,786

PRIOR FILING DATE: 2002-06-26

PRIOR APPLICATION NUMBER: 60/460,787

PRIOR FILING DATE: 2003-04-04

PRIOR APPLICATION NUMBER: 10/606,320

PRIOR FILING DATE: 2003-06-25

NUMBER OF SEQ ID NOS: 134

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 71

LENGTH: 2028

TYPE: DNA

ORGANISM: Bacillus thuringiensis (mutated)

FEATURES:

NAME/KEY: CDS

LOCATION: (1)...(2028)

US-11-058-727-71

Query Match 2.7%; Score 59.6; DB 7; Length 2028;
Best Local Similarity 47.1%; Pred. No. 0.023;

Matches 292; Conservative 0; Mismatches 289; Indels 39; Gaps 2;

367 GACAAAACAGTATGACACAAATTTTAAATGGAGAAATTTTGTGTATACCGTTA 426
322 GAAAAGTCAATGGAGAAATTTTATGAAACAAGTAAAGAACTCATTAATCAAAAAATA 381
427 ACAGAAAGCATTAACAGCTAAAGTTACAACTTTAGAGAGTTTACCAATTTACAA 486
382 GCAGAAATATGCAAGAAATTAAGCGCTTTCGAAATTTAGAAAGATTAGTAAATTAACCA 441
487 AGCTAATATACAGCATTAATGATGATTTGAGAAATTTAAAGATCAAGCTCCGTGATTA 546
442 TTATATCTAATCGCTGCTTGAAGATGGAAGAAATTCATTTCCAGTCGAGGTTTTCGA 501
547 CCACCATCATGACATTAACAAGCTGCTTGAATCTTAAATTAACAATTTGAGAAATGTT 606
502 AGTGAAGGTCAAGCTTACGAGA-----TGTGCAAAATCGATTTGAATCCTG 549
607 CACAAATGTTTATTTGCAAGAAATACCTGTTCCAACTTGAATCTTAATAACGCTATTA 666
550 GATAGTTTATTTACGCAATATATGCAATCTTTAGAGTGAACAAATTTGAAGTACCAATTC 609
667 CTACCTATTTATGCGCAAGCTGCTAATTTTCAATTTAATTTATTAACAAGAGTCTGAA 726
610 CTACTGTTATATGCAATGCGCAAGCTTCACTTATTTACTTTATTAAGAGCGCTCAATT 669
727 TTGCTGATGATGAGATGCAATATACATCTTCAAAATTTGAACCTTAATGCTGGAACA 786
670 TTGGAAGAGATGGGATGTCACAACT-----ACT 702
787 TCAGATGATCTATTAATCTTTAAAGAAATATACCTAAATATAGTAATCTATTTGCA 846


```
Db 382 GCAGAAATATGCAAGAAATAAAGCGCTTTCGAAATTGAGAGATTAAGTAATAATTACCAA 441
Qy 487 AGCTATATATACAGATTAGATGATTTGGAGAAAATTTAAAAGACTACAAAGCTCTGATTA 546
Db 442 TTATATCTAACTGGCTTGAAGAAATGGAGAGAAAATCAATTTCGAGTCGAGTTTTCGA 501
Qy 547 CCACATCATCAGCATTTACAACAAGCTGCTTGAAGCTTTAAATTCAGATTGAGAAATGTT 606
Db 502 AGTCAGAGTCCAGCTTACGAGA-----TGTGCGAAAATTCGATTTGAATCTCTG 549
Qy 607 CACAATGATTTATTCGAGAAATACCTGTTTCCAACTGTAACCTTAATTAAGCTATTA 666
Db 550 GATGCTTATTTATGCGCAATATATGCGCATCTTTTGAAGTACAAAATTTGAAGTACCAATTC 609
Qy 667 CTACCTATTTATGCGCAAGCTGCTTAATTTTCATTTAAATTTTATTAACAACAAGTCTGAA 726
Db 610 CTTACTGATATGCAATGGGAGCCCACTTCATTTACGTTATTTAAAGAGCGGCTCAATT 669
Qy 727 TTGCTGATGATGCAATGCAATATACATCTTCAAAATTTGAACCTTAATGCTGGAACA 786
Db 670 TTTGAGAGAAATGCGGATGCTCAACACT-----ACT 702
Qy 787 TCAGATGCTATTTATTAACCTTTTAAAGAAATATACCTAATATATATTAATTAATTTGCA 846
Db 703 ATTAATATCTATATGATGTCGCAATGAACTTACTGCAAGATATTCGATCAGCTGTGTA 762
Qy 847 AATACCTATAGACAGCACTAAAAATCTTAGACAGCAACAATATATGAAATGAGATTA 906
Db 763 AAGTGTATGAAATCGTTTGAAGCAAAATTAAGGACAGACGGCTTAACATGAGTTGAC 822
Qy 907 TTTAATGACTATCGAAGATATATGACCACTTAATGATTAATGATTAATGATTAATTTCT 966
Db 823 TATAACCAATTCCTGAGAAATGACACTGCGGTTTATGATGTTGTCATTTATCCCA 882
Qy 967 TTATATGATATATAAAGATA 986
Db 883 AATTATGACACGACGACGTA 902

RESULT 12
US-11-058-727-13
; Sequence 13, Application US/11058727
; Publication No. US20050261483A1
; GENERAL INFORMATION:
; APPLICANT: Andre R. Abad
; APPLICANT: Ronald D. Flannagan
; APPLICANT: Rafael Herrmann
; APPLICANT: Theodore W. Kahn
; APPLICANT: Albert L. Lu
; APPLICANT: Billy Fred McCutchen
; APPLICANT: James K. Presnall
; APPLICANT: James F.H. Wong
; APPLICANT: Cao-Guo Yu
; TITLE OF INVENTION: Gene Encoding Proteins With Pesticidal
; TITLE OF INVENTION: Activity
; FILE REFERENCE: 35718/287809
; CURRENT APPLICATION NUMBER: US/11/058,727
; PRIOR FILING DATE: 2005-02-15
; PRIOR APPLICATION NUMBER: 60/391,786
; PRIOR FILING DATE: 2002-06-26
; PRIOR APPLICATION NUMBER: 60/460,787
; PRIOR FILING DATE: 2003-04-04
; PRIOR APPLICATION NUMBER: 10/606,320
; PRIOR FILING DATE: 2003-06-25
; NUMBER OF SEQ ID NOS: 134
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13
; LENGTH: 2022
; TYPE: DNA
; ORGANISM: Bacillus thuringiensis (truncated)
; FEATURE:
; NAME/KEY: CDS
```

```
; LOCATION: (1)...(2022)
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (0)...(0)
; OTHER INFORMATION: 1218-2A
US-11-058-727-13

Query Match 2.6%; Score 58.2; DB 7; Length 2022;
Best Local Similarity 49.9%; Pred. No. 0.041;
Matches 199; Conservative 0; Mismatches 173; Indels 27; Gaps 1;

Qy 588 AATACGATTTGAGAAATTTGACAAATGATTTTATTCGAGAAATACCTGTTTCCAACTTGA 647
Db 510 AATGCAATTTGAATCCGTGATAGTTTATTTATGCAATATACATGCAATCTTTTGCAGTGAC 569
Qy 648 AACTTATTAACGCTATTTACTATTTTATGCGCAAGCTGCTAATTTTCAATTTAAATTT 707
Db 570 AATTTTGAAGTACATCTCTTACAGATATATACAGACGCGCAACTTCATTTACTGTT 629
Qy 708 ATTAACAAGAGTGTGAATTTGCTGATGATGAAATGCAAGATATACATCTTCACAAT 767
Db 630 ATTAAGAGCGCTTCAATTTTGGAGAAATGG----- 663
Qy 768 TGAACCTAATGCTGGAACATCAGATGACTATTAATTTTAAAGAAATATACCTTA 827
Db 664 -GATGTGCTTACACCACTATTAATTAATTAATGATGTCGCAATGAACTTACTGACA 722
Qy 828 ATATGATATATTTGTCAAATACCTATTAAGACAGCACTAAAAATCTTAGAGAGAAC 887
Db 723 ATATTCGATCACTGTGTAAAGGTATGAACTGTTTACGAAATTAAGGACAGAG 782
Qy 888 AATATGAAGTGAAGTATTTAATGACATTCGAGATATATGACCACTTAATCTGATTTAGA 947
Db 783 CGCTAAACAATGGGTGACATTAACCAATTCCTGAGAAATGACATGACGCTTTTGA 842
Qy 948 TACATCTTCAATTTTCTTTATATGATATATAAAGATA 986
Db 843 TGTGTTCATTTATCCCAATATATGACACGACGTA 881

RESULT 13
US-11-108-389-13
; Sequence 13, Application US/11108389
; Publication No. US20050261188A1
; GENERAL INFORMATION:
; APPLICANT: Andre R. Abad
; APPLICANT: Ronald D. Flannagan
; APPLICANT: Rafael Herrmann
; APPLICANT: Theodore W. Kahn
; APPLICANT: Albert L. Lu
; APPLICANT: Billy Fred McCutchen
; APPLICANT: James K. Presnall
; APPLICANT: James F.H. Wong
; APPLICANT: Cao-Guo Yu
; TITLE OF INVENTION: Gene Encoding Proteins With Pesticidal
; TITLE OF INVENTION: Activity
; FILE REFERENCE: 35718/291049
; CURRENT APPLICATION NUMBER: US/11/108,389
; PRIOR FILING DATE: 2005-04-18
; PRIOR APPLICATION NUMBER: 60/391,786
; PRIOR FILING DATE: 2002-06-26
; PRIOR APPLICATION NUMBER: 60/460,787
; PRIOR FILING DATE: 2003-04-04
; PRIOR APPLICATION NUMBER: 10/606,320
; PRIOR FILING DATE: 2003-06-25
; NUMBER OF SEQ ID NOS: 134
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13
; LENGTH: 2022
; TYPE: DNA
; ORGANISM: Bacillus thuringiensis (truncated)
; FEATURE:
; NAME/KEY: CDS
```

LOCATION: (1)...(2022)
FEATURE:
NAME/KEY: misc feature
LOCATION: (0)...(0)
OTHER INFORMATION: 1218-2A
US-11-108-389-13

Query Match 2.6%; Score 58.2; DB 7; Length 2022;
Best Local Similarity 49.9%; Pred. No. 0.041;
Matches 199; Conservative 0; Mismatches 173; Indels 27; Gaps 1;

OY 588 AATACGATTTGAGAAATGTTCAATGATTTTATTCGAGAAATACCTGGTTCCAACTTGA 647
DB 510 AATCGATTTGAAATCCGATGATTTTATTAACCAATACATGCGATCTTTTCAGTGAC 569
OY 648 AACTATTAACGCTATTAACCTATTTATGCGAAGCTGTGTAATTTTCAATTAATTT 707
DB 570 AATTTTGAAGTACCATTCCTTAACGATATTAACAGACGACCACTTCAATTTACTGTT 629
OY 708 ATTACAAAGAGTCTGATTTGCTGATGAAATGGAATGCAATATATACATCTTCACAAT 767
DB 630 ATTAAAGGACGCTTCAATTTTGGAGAAAGATG----- 663
OY 768 TGAACCTAATGCTGGAACATGATGATTTATTAACCTTTTAAAGAAATATACCTPA 827
DB 664 -GATGCTCAACACCACTATTAATTAATGATGATGCTCAATGAAATGAACTTACTGACA 722
OY 828 ATATAGTAACTATTTGCAAAATACCTATTAAGAACAGACTTAAATACTTAAAGACGAA 887
DB 723 ATATTTGATACCTGTGTAAAGTGTATGAACCTGTTTAAAGAAATTAAGGACGAG 782
OY 888 AATATGAATGAGATATATTAATGACTATGAAATGATATGACATTAATCTGATTTAGA 947
DB 783 CGCTAAACATGGGTCGACTATTAACCAATTCCTGTAGAAATGACACTGACGTTTTAGA 842
OY 948 TACCATCTCTCAATTTTCTTTATATGATATTAAGAAATAT 986
DB 843 TGTGTGCTATTAATCCCAATATTAAGACACGACGTA 881

RESULT 14
US-11-058-727-3
Sequence 3, Application US/11058727
Publication No. US20050261483A1
GENERAL INFORMATION:
APPLICANT: Andre R. Abad
APPLICANT: Ronald D. Flannagan
APPLICANT: Rafael Herimann
APPLICANT: Theodore W. Kahn
APPLICANT: Albert L. Lu
APPLICANT: Billy Fred McCutchen
APPLICANT: James K. Presnail
APPLICANT: James F.H. Wong
APPLICANT: Cao-Guo Yu
TITLE OF INVENTION: Genes Encoding Proteins With Pesticidal
TITLE OF INVENTION: Activity
FILE REFERENCE: 35718/287809
CURRENT APPLICATION NUMBER: US/11/058,727
CURRENT FILING DATE: 2005-02-15
PRIOR APPLICATION NUMBER: 60/391,786
PRIOR FILING DATE: 2002-06-26
PRIOR APPLICATION NUMBER: 60/460,787
PRIOR FILING DATE: 2003-04-04
PRIOR APPLICATION NUMBER: 10/606,320
PRIOR FILING DATE: 2003-06-25
NUMBER OF SEQ ID NOS: 134
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 3
LENGTH: 3633
TYPE: DNA
ORGANISM: Bacillus thuringiensis
FEATURE:
NAME/KEY: CDS

LOCATION: (1)...(3633)
FEATURE:
NAME/KEY: misc feature
LOCATION: (0)...(0)
OTHER INFORMATION: CRY1218-2
US-11-058-727-3

Query Match 2.6%; Score 58.2; DB 7; Length 3633;
Best Local Similarity 49.9%; Pred. No. 0.046;
Matches 199; Conservative 0; Mismatches 173; Indels 27; Gaps 1;

OY 588 AATACGATTTGAGAAATGTTCAATGATTTTATTCGAGAAATACCTGGTTCCAACTTGA 647
DB 510 AATCGATTTGAAATCCGATGATTTTATTAACCAATACATGCGATCTTTTCAGTGAC 569
OY 648 AACTATTAACGCTATTAACCTATTTATGCGAAGCTGTGTAATTTTCAATTAATTT 707
DB 570 AATTTTGAAGTACCATTCCTTAACGATATTAACAGACGACCACTTCAATTTACTGTT 629
OY 708 ATTACAAAGAGTCTGATTTGCTGATGAAATGGAATGCAATATATACATCTTCACAAT 767
DB 630 ATTAAAGGACGCTTCAATTTTGGAGAAAGATG----- 663
OY 768 TGAACCTAATGCTGGAACATGATGATTTATTAACCTTTTAAAGAAATATACCTPA 827
DB 664 -GATGCTCAACACCACTATTAATTAATGATGATGCTCAATGAAATGAACTTACTGACA 722
OY 828 ATATAGTAACTATTTGCAAAATACCTATTAAGAACAGACTTAAATACTTAAAGACGAA 887
DB 723 ATATTTGATACCTGTGTAAAGTGTATGAACCTGTTTAAAGAAATTAAGGACGAG 782
OY 888 AATATGAATGAGATATATTAATGACTATGAAATGATATGACATTAATCTGATTTAGA 947
DB 783 CGCTAAACATGGGTCGACTATTAACCAATTCCTGTAGAAATGACACTGACGTTTTAGA 842
OY 948 TACCATCTCTCAATTTTCTTTATATGATATTAAGAAATAT 986
DB 843 TGTGTGCTATTAATCCCAATATTAAGACACGACGTA 881

RESULT 15
US-11-108-389-3
Sequence 3, Application US/11108389
Publication No. US20050261188A1
GENERAL INFORMATION:
APPLICANT: Andre R. Abad
APPLICANT: Ronald D. Flannagan
APPLICANT: Rafael Herimann
APPLICANT: Theodore W. Kahn
APPLICANT: Albert L. Lu
APPLICANT: Billy Fred McCutchen
APPLICANT: James K. Presnail
APPLICANT: James F.H. Wong
APPLICANT: Cao-Guo Yu
TITLE OF INVENTION: Genes Encoding Proteins With Pesticidal
TITLE OF INVENTION: Activity
FILE REFERENCE: 35718/291049
CURRENT APPLICATION NUMBER: US/11/108,389
CURRENT FILING DATE: 2005-04-18
PRIOR APPLICATION NUMBER: 60/391,786
PRIOR FILING DATE: 2002-06-26
PRIOR APPLICATION NUMBER: 60/460,787
PRIOR FILING DATE: 2003-04-04
PRIOR APPLICATION NUMBER: 10/606,320
PRIOR FILING DATE: 2003-06-25
NUMBER OF SEQ ID NOS: 134
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 3
LENGTH: 3633
TYPE: DNA
ORGANISM: Bacillus thuringiensis
FEATURE:
NAME/KEY: CDS

```
; LOCATION: (1)...(3633)
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (0)...(0)
; OTHER INFORMATION: CY1218-2
US-11-108-389-3
```

```
Query Match      2.6%; Score 58.2; DB 7; Length 3633;
Best Local Similarity 49.9%; Pred. No. 0.046;
Matches 199; Conservative 0; Mismatches 173; Indels 27; Gaps 1;
```

```
QY 588 AATGAGATTGAGATGTTCAATGATTTATTCGAGAAATACCTGTTCCAACTTGA 647
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 510 AATGAGATTGAGATGTTCAATGATTTATTCGAGAAATACCTGTTCCAACTTGA 647
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 648 AACTTATTAACGCTATTACTACCTATTATTCGCAAGCTGCTAATTTTCATTAAATTT 707
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 570 AATTTGAGATGATCATTCTTACGATATATACAGGACCAACCTTCATTACTGTT 629
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 708 ATTACAAACAGGTGCTGATTTGGCTGATGAATGGAATGAGATATACATCCTTCACAAAT 767
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 630 ATTAAAGGACCGCTTCAATTTTGGAGAAATGG----- 663
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 768 TGAACCTTAATGCTGGAACATGAGATGATTAATTAATTAAGAAATATACCTTA 827
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 664 -GGAATGCTTAACAAACCACTATTATTAATGATGTCAAATGAAATGAACTTACTGAGA 722
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 828 ATATAGTAATCTATTGTGCAAAATACCTATAGAAACAGACTAAATAATCTTAGAGACGAAC 887
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 723 ATATTCTGATCAGCTGTGTAAGGTGTATGAAAGTGTATTACCAAAATTAAGGACAGAG 782
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 888 AAATATGAAATGAGATATATTAATGACTATGAGATATATGACCATTTACTGTATTAGA 947
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 783 CGCTAAACAAATGGGTGCACTATTAACCAATTCGTTAGAAATGACACTGACGTTTGA 842
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 948 TACCATCTTCATTTCTTTATATGATTAATAAAGATA 986
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 843 TGTGTGTGCAATTATTCCTCAAAATATAGACACGACGTA 881
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

Search completed: December 11, 2005, 14:11:53
Job time : 223 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using SW model

Run on: December 11, 2005, 03:43:39 ; Search time 1791 Seconds

(without alignments)
10194.744 Million cell updates/sec

Title: US-10-783-417-1

Perfect score: 2208
Sequence: 1 atgatacaataacagataa.....atcacataaacacttga 2208

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 9793542 seqs, 4134689005 residues

Total number of hits satisfying chosen parameters: 19587084

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications NA_Main:
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
3: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:*
4: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*
5: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
6: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
7: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
8: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq:*
9: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq:*
10: /cgn2_6/ptodata/1/pubpna/US11_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2208	100.0	2208	8	US-10-783-417-1
2	1903.8	86.2	2235	8	US-10-782-570-1
3	1794.4	81.3	2085	8	US-10-782-570-3
4	163.8	7.4	4896	6	US-09-756-526A-3
5	163.8	7.4	4896	6	US-10-345-020-3
6	163.8	7.4	4896	6	US-10-342-821-3
7	98.4	4.5	3504	5	US-10-089-678-2
8	98.4	4.5	3690	5	US-10-089-678-3
9	97	4.4	3673778	6	US-10-312-841-1
10	94.6	4.1	15548	6	US-10-311-455-2128
11	90.4	4.1	8056	6	US-10-473-126-386
12	85.8	3.9	3522	3	US-09-826-660-5
13	85.8	3.9	3522	3	US-09-837-961-7
14	85.8	3.9	3522	8	US-10-825-751-7
15	81.2	3.7	6175	6	US-10-311-455-1880
16	77.8	3.5	3673778	6	US-10-312-841-2
17	76.4	3.5	3507	6	US-10-428-961-21
18	74.2	3.4	8056	6	US-10-473-126-240
19	73.2	3.3	1959	7	US-10-614-076-13
20	72.6	3.3	6641	6	US-10-311-455-288
21	72.6	3.3	6641	6	US-10-240-452-36
22	72.6	3.3	8237	7	US-10-221-714A-527
23	71.8	3.3	7498	6	US-10-311-455-230

24	71.6	3.2	1959	7	US-10-614-076-27	Sequence 27, Appl
25	71.6	3.2	3684	8	US-10-929-754-2	Sequence 2, Appl
26	71	3.2	8056	9	US-10-473-126-386	Sequence 386, Appl
27	70.8	3.2	11092	6	US-10-311-455-1486	Sequence 1486, Ap
28	70.4	3.2	5928	6	US-10-311-455-2059	Sequence 2059, Ap
29	70.4	3.2	11745	6	US-10-240-453-206	Sequence 206, App
30	70	3.2	1482	7	US-10-614-076-69	Sequence 69, Appl
31	70	3.2	1956	7	US-10-614-076-51	Sequence 51, Appl
32	70	3.2	1956	7	US-10-614-076-55	Sequence 55, Appl
33	70	3.2	1956	7	US-10-614-076-57	Sequence 57, Appl
34	70	3.2	1959	5	US-10-232-665-1	Sequence 1, Appl
35	70	3.2	1959	7	US-10-614-076-1	Sequence 19, Appl
36	70	3.2	1959	7	US-10-614-076-19	Sequence 41, Appl
37	70	3.2	1959	7	US-10-614-076-41	Sequence 43, Appl
38	70	3.2	1959	7	US-10-614-076-43	Sequence 45, Appl
39	70	3.2	1959	7	US-10-614-076-45	Sequence 65, Appl
40	70	3.2	1959	7	US-10-614-076-65	Sequence 67, Appl
41	70	3.2	1959	7	US-10-614-076-67	Sequence 97, Appl
42	70	3.2	1959	7	US-10-614-076-97	Sequence 102, App
43	70	3.2	2280	7	US-10-614-076-102	Sequence 3, Appl
44	70	3.2	3507	7	US-10-614-524-3	Sequence 17, Appl
45	69.4	3.1	1959	7	US-10-614-076-17	

ALIGNMENTS

RESULT 1
US-10-783-417-1
Sequence 1, Application US/10783417
Publication No. US20040216186a1
GENERAL INFORMATION:
APPLICANT: Carozzi, Nadine
APPLICANT: Kozziel, Michael G.
APPLICANT: Duck, Nicholas B.
APPLICANT: Carr, Brian
TITLE OF INVENTION: AXMI-006, A Delta-Endotoxin Gene and
TITLE OF INVENTION: Methods for Its Use
FILE REFERENCE: 045600/274146
CURRENT APPLICATION NUMBER: US/10/783,417
CURRENT FILING DATE: 2004-02-20
PRIOR APPLICATION NUMBER: 60/448,806
PRIOR FILING DATE: 2003-02-20
NUMBER OF SEQ ID NOS: 15
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 1
LENGTH: 2208
TYPE: DNA
ORGANISM: Bacillus thuringiensis
FEATURE:
NAME/KEY: CDS
LOCATION: (1)...(2208)
US-10-783-417-1

Query Match 100.0%; Score 2208; DB 8; Length 2208;
Best local similarity 100.0%; Pred. No. 0;
Matches 2208; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	ATGAATCAAAATTAACGATATACGAAATATGATATGATTCGATACCTTACCTTAT	60
DB	1	ATGAATCAAAATTAACGATATACGAAATATGATATGATTCGATACCTTACCTTAT	60
QY	61	TTTCGAAACAAACAGTAATGATTCGATACCTTACGAAATATGATTCGATACCTTAT	120
DB	61	TTTCGAAACAAACAGTAATGATTCGATACCTTACGAAATATGATTCGATACCTTAT	120
QY	121	TTTCAAAACAAATTAACGATATGATTCGATACCTTACGAAATATGATTCGATACCTTAT	180
DB	121	TTTCAAAACAAATTAACGATATGATTCGATACCTTACGAAATATGATTCGATACCTTAT	180
QY	181	GATATTTGAGACATTTGCTGATGCTGATATGCTGATGCTGATGCTGATGCTGATGCTGAT	240

Dh 181 GATAATTTGAGACATTTGCTAGTGTGATACAAATGCTGCAAGTTAGTGCAGGTAATTT 240
Qy 241 GATTCGGGTACTGTTTAGCCGGTATAGGTGGGCTCACTTGTATATCCGAGCCGATAGGA 300
Dh 241 GATTCGGGTACTGTTTAGCCGGTATAGGTGGGCTCACTTGTATATCCGAGCCGATAGGA 300
Qy 301 ATATAGGTGCTATATATATATCTTTTGGTACCCCTAATCACTGTCTTTTGGCCCGCGGA 360
Dh 301 ATATAGGTGCTATATATATCTTTTGGTACCCCTAATCACTGTCTTTTGGCCCGCGGA 360
Qy 361 GAACAAGCAAAAAGATAGACACAATTTATTAATGGAGAAATTTTGTGTATCA 420
Dh 361 GAACAAGCAAAAAGATAGACACAATTTATTAATGGAGAAATTTTGTGTATCA 420
Qy 421 CCGTTAACAGAAAGATATAACAGTTAAAGTTACAACTTTAGAAAGATTAGCAATA 480
Dh 421 CCGTTAACAGAAAGATATAACAGTTAAAGTTACAACTTTAGAAAGATTAGCAATA 480
Qy 481 TTACAAGCTATATATACAGATTAGATAGTGGAGAAAATTAATAAGCTACAGCTCCT 540
Dh 481 TTACAAGCTATATATACAGATTAGATAGTGGAGAAAATTAATAAGCTACAGCTCCT 540
Qy 541 GGATTTACACCATCATCGACATTACACAGAGCTGCTTGACTCTTAATATCGATTGAG 600
Dh 541 GGATTTACACCATCATCGACATTACACAGAGCTGCTTGACTCTTAATATCGATTGAG 600
Qy 601 AATGTTCAATGATTTTATTGAGAAATACCTGGTTCCAACTTGAACCTTATATAACG 660
Dh 601 AATGTTCAATGATTTTATTGAGAAATACCTGGTTCCAACTTGAACCTTATATAACG 660
Qy 661 CTATTACTACCTATTTATGCGCAAGCTGCTAATTTTCTTTAAATTTTATACCAAGCT 720
Dh 661 CTATTACTACCTATTTATGCGCAAGCTGCTAATTTTCTTTAAATTTTATACCAAGCT 720
Qy 721 GCTGAATTTGGCTGATGGAATGGAATGAGATATACATCTTCAACAATTTGAACCTATGCT 780
Dh 721 GCTGAATTTGGCTGATGGAATGGAATGAGATATACATCTTCAACAATTTGAACCTATGCT 780
Qy 781 GGAACATCAGATGATCTATTATTAACCTTTAAAGAAAATATACCTAATATATGTAATCT 840
Dh 781 GGAACATCAGATGATCTATTATTAACCTTTAAAGAAAATATACCTAATATATGTAATCT 840
Qy 841 TGTCGCAATACCTATATAGAACAGACCTAATAATCTTAGAGACGAACCAATATGAAATGG 900
Dh 841 TGTCGCAATACCTATATAGAACAGACCTAATAATCTTAGAGACGAACCAATATGAAATGG 900
Qy 901 AGTATATTTATGATCTATCGAAGATATATGACATCTGATTTAGATACATCTTCA 960
Dh 901 AGTATATTTATGATCTATCGAAGATATATGACATCTGATTTAGATACATCTTCA 960
Qy 961 TTTTCTTTATATGATATTAATAAGATATAGAGATTCAATAGAGAAATAGAAATGAAAGGC 1020
Dh 961 TTTTCTTTATATGATATTAATAAGATATAGAGATTCAATAGAGAAATAGAAATGAAAGGC 1020
Qy 1021 ATTAAGAAATGAACTCAAGAGAAATTTATACATGAAATTAATTTTGTGCTTCT 1080
Dh 1021 ATTAAGAAATGAACTCAAGAGAAATTTATACATGAAATTAATTTTGTGCTTCT 1080
Qy 1081 CAACCTAGAGTTCAACCAATCTAGCTAGATGGAATTAATTTTAAACGCGGAAGTTT 1140
Dh 1081 CAACCTAGAGTTCAACCAATCTAGCTAGATGGAATTAATTTTAAACGCGGAAGTTT 1140
Qy 1141 AAATATATTTTCAATTTTGAACCAATTTATTTTATACAGAAAATTAACAAATTTGGGAAAT 1200
Dh 1141 AAATATATTTTCAATTTTGAACCAATTTATTTTATACAGAAAATTAACAAATTTGGGAAAT 1200
Qy 1201 CGTTAGTGTATTTCTATTCGATGACACTTATAGCAATCTATTAACCTGAAACT 1260
Dh 1201 CGTTAGTGTATTTCTATTCGATGACACTTATTAACCTTATTAACCTGAAACT 1260
Qy 1261 TTATATGAGAAAGAAAGAGTTGACCCGACAAACAAATTAAGCAATTTGAATCTTAT 1320
Dh 1261 TTATATGAGAAAGAAAGAGTTGACCCGACAAACAAATTAAGCAATTTGAATCTTAT 1320

Qy 1321 AAAGTTTCAATTTGTAATGATAGCAATCACTCTGTTTTCCCTATTTCAACCACTTT 1380
Dh 1321 AAAGTTTCAATTTGTAATGATAGCAATCACTCTCTGTTTTCCCTATTTCAACCACTTT 1380
Qy 1381 ATATATTAATCAAAATTTGAACCTTTATTTAAATGAGCTCATCTTAACAACACTCAATATTTCA 1440
Dh 1381 ATATATTAATCAAAATTTGAACCTTTATTTAAATGAGCTCATCTTAACAACACTCAATATTTCA 1440
Qy 1441 GCAGAGAGGTCTTTATCTAATTTATCAAAACCAACTTTTTCATTTCTAGAAAAAA 1500
Dh 1441 GCAGAGAGGTCTTTATCTAATTTATCAAAACCAACTTTTTCATTTCTAGAAAAAA 1500
Qy 1501 GACTGCAATCTAGTTATGATTCAGGTTGTTCAACCAACTTTTATTAATCTATAGCATATT 1560
Dh 1501 GACTGCAATCTAGTTATGATTCAGGTTGTTCAACCAACTTTTATTAATCTATAGCATATT 1560
Qy 1561 TTATCCATTTTTCATTTATTTACTTATTTCTATGATGATGGAATTAAGCTACAAATTTTA 1620
Dh 1561 TTATCCATTTTTCATTTATTTACTTATTTCTATGATGATGGAATTAAGCTACAAATTTTA 1620
Qy 1621 GATACAGGTGATTTAGATGGAACAACAGATGTTGATAGATTAATGCAATATCAGAT 1680
Dh 1621 GATACAGGTGATTTAGATGGAACAACAGATGTTGATAGATTAATGCAATATCAGAT 1680
Qy 1681 AAAATTAATCAATGATCCAGCAATCAAGGTTAACAATCTTGATACAACTTAAGATA 1740
Dh 1681 AAAATTAATCAATGATCCAGCAATCAAGGTTAACAATCTTGATACAACTTAAGATA 1740
Qy 1741 ATTGAAGAGCTGCTGATACAGAGAGAAATCTGTTTATTTACAAATCAAGGCGTTTA 1800
Dh 1741 ATTGAAGAGCTGCTGATACAGAGAGAAATCTGTTTATTTACAAATCAAGGCGTTTA 1800
Qy 1801 GAAATTCATGTAATACTCTTAATTTCTACACAATCTTATTTCACTTAGATATGCT 1860
Dh 1801 GAAATTCATGTAATACTCTTAATTTCTACACAATCTTATTTCACTTAGATATGCT 1860
Qy 1861 ACAATGCTGCTGAAATATCTTCTTAATATCTCTTAATATCTCTTAATATCAAGAGATATAGGA 1920
Dh 1861 ACAATGCTGCTGAAATATCTTCTTAATATCTCTTAATATCTCTTAATATCAAGAGATATAGGA 1920
Qy 1921 ATACCACTTCAAGACCTCAACCAACTTTTCTGATGCAAAATTAATTAATTAATCAATAC 1980
Dh 1921 ATACCACTTCAAGACCTCAACCAACTTTTCTGATGCAAAATTAATTAATTAATCAATAC 1980
Qy 1981 GGAATTTTGGGTATTTCCAAATTTCCAAATGACAGTAACATTTACCTTAATTCGAAACATA 2040
Dh 1981 GGAATTTTGGGTATTTCCAAATTTCCAAATGACAGTAACATTTACCTTAATTCGAAACATA 2040
Qy 2041 CCAATTAATTTAATCGTGAGATGTAACAATTTCAATTTTAATCATTTGATTAATAATGAA 2100
Dh 2041 CCAATTAATTTAATCGTGAGATGTAACAATTTCAATTTTAATCATTTGATTAATAATGAA 2100
Qy 2101 TTTATACCAATTTCTTCTCTATAGCAACCAAAATTAAGAAAAAATTAAGAACTATC 2160
Dh 2101 TTTATACCAATTTCTTCTCTATAGCAACCAAAATTAAGAAAAAATTAAGAACTATC 2160
Qy 2161 CAAACAAAATTAATCAATTTTTCACAAATCATACAAAACACTTTTGA 2208
Dh 2161 CAAACAAAATTAATCAATTTTTCACAAATCATACAAAACACTTTTGA 2208

RESULT 2
US-10-782-570-1
; Sequence 1, Application US/10782570
; Publication No. US20040210965A1
; GENERAL INFORMATION:
; APPLICANT: Carozzi, Nadine
; APPLICANT: Hargiles, Tracy
; APPLICANT: Koziele, Michael G.
; APPLICANT: Duck, Nicholas B.
; APPLICANT: Carr, Brian
; TITLE OF INVENTION: AxiM-007, A Delta-Endotoxin Gene and

Db 1315 ACAACGTGTTTTCATTTCCGTGAATAAAAAAGCTGAACAATAATTAAATCAAAATGGT 1374

Oy 1511 TCACCAAACCTTTAATTAAGTATATGTCATATTTTATGCCATTTTTACTATTTACTATTCC 1590

Db 1375 TTACCAAGCCTTAATTAAGTATATGTCATATTTTATGCCAGTTTCTTTATTTAAATATCC 1434

Oy 1591 TATGATATGGATTAACAGCTAACAAATATTAGTATACAGTGATTAAGATGAGACACAGT 1650

Db 1435 TATTAATATGGATTAAGCCCTAAATATATTATTAATACAGGTGATTAAGATGAGACACAGT 1494

Oy 1651 AGGTGTGATAGATATTAATGCAATATCAGATAAAATTAATTAACAATGATCCAGCAATCAA 1710

Db 1495 AGGTGATTAAGAATTAATGCAATATCAGATAAAATTAATTAACAATGATCCAGCAATCAA 1554

Oy 1711 GGTAACAATCTTTGATCAAACTCTAAGGTAATTGAAGAACCTGGTCATACAGAGGAAC 1770

Db 1555 GGTAAACAGTCTTGATACAAACTCTAAGGTAATTGAAGAACCTGGTCATACAGAGGAAC 1614

Oy 1771 TTGGTTATTTTACAAAGTCAAGGGCGTTTACAAANTTAATGTAAGAACCTCCATATTTTCA 1830

Db 1615 TTGGTTATTTTACAAAGTCAAGGGCGTTTACAAANTTAATGTAAGAACCTCCATATTTTCA 1674

Oy 1831 CAATCTTATTTTCAATAGCTTGATATGCTCAAAATGGTGTGAAATTAATCTTCCATAT 1890

Db 1675 CAATCTTATTTTCAATAGCTTGATATGCTCAAAATGGTGTGAAATTAATCTTCCATAT 1734

Oy 1891 ATATCTCTTAACAATACAGAGTAATAGGAATACCACCTCAACGATCTCAACAACATTTT 1950

Db 1735 ATATCTCTTAACAATACAGAGTAATAGGAATACCACCTCAACGATCTCAACAACATTTT 1794

Oy 1951 TCTGGTACAAATTTAATTAATTTTACAAATACGAGATTTTGGGTATTTCCAATTTCCAGT 2010

Db 1795 TCTGGTACAAATTTAATTAATTTTACAAATACGAGATTTTGGGTATTTCCAATTTCCAGT 1854

Oy 2011 ACAGTAACATTAACCTTTAAATCGAAACATACCATTTATATTAAATCGAGAGATGATCA 2070

Db 1855 ACAGTAACATTAACCTTTAAATCGAAACATACCATTTATATTAAATCGAGAGATGATCA 1914

Oy 2071 AATTCATTTTAAATCATTTGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2130

Db 1915 AATTCATTTTAAATCATTTGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1974

Oy 2131 AATGAGAAAAACAAAAATTGAAACTATCGAAACAAAAATTAATTAATTAATTAATTAAT 2190

Db 1975 AATGAGAAAAACAAAAATTGAAACTATCGAAACAAAAATTAATTAATTAATTAATTAAT 2034

Oy 2191 CATCAAAAAACACTTGA 2208

Db 2035 CATCAAAAAATTAATTTA 2052

RESULT 4
US-09-756-526A-3

Sequence 3, Application US/09756526A

Patent No. US20020038005A1

GENERAL INFORMATION:

APPLICANT: Jana, Wojciechowska

APPLICANT: Evgeniy, Lewitin

APPLICANT: Ludmila, Revina

APPLICANT: Igor, Zalunin

APPLICANT: Galina, Chesukhina

TITLE OF INVENTION: NOVEL DELTA-ENDOTOXINS AND NUCLEIC ACID SEQUENCES CODING THE REFOR

FILE REFERENCE: S-30913A

CURRENT APPLICATION NUMBER: US/09/756,526A

PRIOR FILING DATE: 2001-01-08

PRIOR APPLICATION NUMBER: US 60/175,158

PRIOR FILING DATE: 2000-01-07

NUMBER OF SEQ ID NOS: 5

SOFTWARE: PatentIn version 3.0

SEQ ID NO 3

LENGTH: 4896

TYPE: DNA

ORGANISM: Bacillus thuringiensis

FEATURE: /
NAME/KEY: source /
LOCATION: .(1)..(4896) /
OTHER INFORMATION: Bacillus thuringiensis supsp. finitimus strain VPM B-1161 (taxon /
OTHER INFORMATION: 29337 /
NAME/KEY: CDS /
LOCATION: (1129)..(4458) /
OTHER INFORMATION: product: Cry28Aa1 delta-endotoxin /
PUBLICATION INFORMATION: /
AUTHORS: Wojciechowska, et al. /
TITLE: Two novel delta-endotoxin gene families cry26 and cry28 from Bacillus /
TITLE: thuringiensis ssp. finitimus /
JOURNAL: FEBS Lett. /
VOLUME: 453 /
ISSUE: 12 /
PAGES: 46-48 /
DATE: 1999-06-18 /
DATABASE ENTRY DATE: ____-__-__ /
US-09-756-526A-3

Query Match 7.4%; Score 163.8; DB 3; Length 4896;
Best Local Similarity 55.9%; Pred. No. 1.7e-20;
Matches 405; Conservative 0; Mismatches 272; Indels 48; Gaps 3;

QY 306 AGTGTCTAATATATATCTTTTGGTACCCCTAATCATCTGTCTTTGGCCCGGGAGACA 365
DB 1293 AGGTGTGTACTTATATCATTTGGAACTTGGCTCCGTTCTTTGGCTGTATCCAGAGA 1352
QY 366 AGA-----CAAAAGATATGAGCAACAATTATTAATAAGGAGAAATTTTGTATACACC 422
DB 1353 AGATCAAAAAAATTTGGTGCACATTTTGAAGACGAGAGAACCTTTTAATCAAC 1412
QY 423 GTTAAAGAAAGCATTAACAGCTAAAGTTACAACTTTGAAGAGATTAGACAAATAT 482
DB 1413 AATTCTAAGCTGTAAAGAAATAGCATTAAGCTCATCTAAATGTTTAAAGATATAT 1472
QY 483 ACAAAAGCTTAATACGCACTTAGATATTTGAGAAATTAAGAAAGTACAGCTCTGG 542
DB 1473 AACGTACTATGAAGAGCAATTAATGATTGGAAGAAATCCAAAGTCAAAATATCTGCAG 1532
QY 543 ATTACACCATCATGACATTAACAAGCTGCTTGACTTTAAATACGATTGAGAA 602
DB 1533 ATTGGTATACAGA-----GATTTGAAAA 1556
QY 603 TGTTCACATGATTTTATTCGAAAAATCCTGGTTCCAACTTGAACTTATAAACGCT 662
DB 1557 CGCTCATTTCAATTTTGTAAAGCAATATGCAACAACCTCACTTCCAGTATGACAT 1616
QY 663 ATTACTACTATTTATGCGCAAGCTGCTAATTTTCATTTAATTTATTAACAAGGTGC 722
DB 1617 ATTAATTAAGTGTCTATACAGAAAGCTCAAAATTTACATTTGAATTTATTAACATCAAGGTG 1676
QY 723 TGAATTTGGTGTGAATGGAATGCAATATACATCTTCACAAATGGAACCTAATGCTGG 782
DB 1677 ACAATTCGGGATCATGGAATGCAATGCAACACATTCACCAAT-----GTTGAA 1727
QY 783 AACATCAAGTGAATTTATAAATCTTTAAAAAGAAATATATCTTAATATAGTAACTAT 842
DB 1728 GTATCAAGGTACTTATTAATGACGAGCTATTTGGTATATATTTGAAAGATATTAATTA 1787
QY 843 TGAATAATCTATAGAACAGACTAAATAATCTTGAAGCAACCAATATGAATGAG 902
DB 1788 CACCAAGACATACATTAAGATTTGAATCACTTAAAGATTAAGAAAAATCACTGGGA 1847
QY 903 TATATTTAATGATATCGAAGATATATGACATTAATCTGTATTAATACATCTCTCAT 962
DB 1848 TGCTTATTAACATATCGTGAAGAAATGACCTTAATTTGATTTGGATCTTGTGCAACT 1907
QY 963 TTCTTTAATGATATTAATAAGATATAGAGTTCAATAGAGAAATAGAAATGAAGCAT 1022
DB 1908 TCCTTTTATGATATACGTGTTTTTCCAGAGAGATGAATTAACAAGAGAGCT 1967
QY 1023 TAAGA 1027

Db 1968 TTATA 1972

RESULT 5

US-10-345-020-3
Sequence 3, Application US/10345020
Publication No. US20030150018A1
GENERAL INFORMATION:
APPLICANT: Jana, Wojciechowska
APPLICANT: Evgeny, Lewitin
APPLICANT: Ludmila, Revina
APPLICANT: Igor, Zalunin
APPLICANT: Galina, Cheslukhina
TITLE OF INVENTION: NOVEL DELTA-ENDOTOXINS AND NUCLEIC ACID SEQUENCES CODING THEREFOR
FILE REFERENCE: S-30913B
CURRENT APPLICATION NUMBER: US/10/345,020
CURRENT FILING DATE: 2003-01-15
PRIOR APPLICATION NUMBER: US 60/175,158
PRIOR FILING DATE: 2000-01-07
NUMBER OF SEQ ID NOS: 5
SOFTWARE: PatentIn version 3.0
SEQ ID NO 3
LENGTH: 4896
TYPE: DNA
ORGANISM: *Bacillus thuringiensis*
FEATURE:
NAME/KEY: source
LOCATION: (1)..(4896)
OTHER INFORMATION: *Bacillus thuringiensis* supsp. *finitimus* strain VKPM B-1161 (taxon
FEATURE:
NAME/KEY: CDS
LOCATION: (1129)..(4458)
OTHER INFORMATION: product: Cry28a1 delta-endotoxin
PUBLICATION INFORMATION:
AUTHORS: Wojciechowska, et al.
TITLE: Two novel delta-endotoxin gene families cry26 and cry28 from *Bacillus*
TITLE: *thuringiensis* ssp. *finitimus*
JOURNAL: *FEBS Lett.*
VOLUME: 453
ISSUE: 12
PAGES: 46-48
DATE: 1999-06-18
US-10-345-020-3

Query Match Best Local Similarity 55.9%; Pred. No. 1.7e-20;

Matches 405; Conservative 0; Mismatches 272; Indels 48; Gaps 3;

QY 306 AGGTGCTAATAATATCTTTGTACCTATATCACTGCTTTTGGCCCGGGAGAAC 365
DB 1293 AGGTGTGCTAATAATATCTTTGTACCTATATCACTGCTTTTGGCCCGGGAGAAC 1352
QY 366 AGA---CAAACAGATGACACAACTTATTAATGGGAAATTTTGTGTACACC 422
DB 1353 AGATCCAAAAAATTTGTGTACAACTTATTAATGGGAAATGGGAAATGGGAAATTTTGTGTACAAAC 1412
QY 423 GTTAACAGAAAGCATAAAGCTAAAGTTACAACTTTAGAGATTGACAAATATT 482
DB 1413 AATTTCTACAGCTGTAAAGAAATAGCATTTAGCTCATTAATGTTTAAAGATGATTT 1472
QY 483 ACAAGCTATATATACAGCATTTAGATTTGAGGAAATTTAAAGATCTACAGCTCTGG 542
DB 1473 AACGTAATGAAAGATTAATGATTTGAGGAAATTCAGATGCAATATCTGCAG 1532
QY 543 ATTACCAATCATATGATTAACAACAAGCTGCTTGACTCTTAATAATGATTTGAGAA 602
DB 1533 ATTGTATCACAGA-----GATTTGAAAA 1556
QY 603 TGTTCACATGATTTTATTTGAGAAATACCTGTTTCCATGTAATTTAAAGAGCT 662
DB 1557 CGCTCATTTCAATTTTGTAGCAATATGCAACTCCAACTTCCACGATATGACACTT 1616

QY 663 ATTACTACTATTTATGCGAAGCTGCTAATTTTCAATTTAATTTATTAACAAGGTGC 722
DB 1617 ATTATTAAGTTGCTATACAGAGCTGCAGAAATTTACATTTGAAATTTATTAACATCAAGGTGT 1676
QY 723 TGAATGGCTGATGAATGGAATGCAATATACATCCCTCCCAATTAACCTAATGCTGG 782
DB 1677 ACAATTCGCGATCAATGAATGCAATCAACCATTTCCCAAT-----GTTGAA 1727
QY 783 AACATGATGACTATTTAATACTTTTAAAGAAATATATCTTAATATAGTACTATTG 842
DB 1728 GTGATGAGTACTATTTATGACGATTTGATTTATTTGAAAGATATTTATTTTGTG 1787
QY 843 TGCATAATCTTATGAACAGAGCTAAAAATCTTTAGAGCAACCAATATGAATGAG 902
DB 1788 CACCAAGACATACATTAAGGATTTGATTCACCTTAAAGAAATCAAGAAAAATCAATGGA 1847
QY 903 TATATTAATGATGAGATATATATGACATTTCTGTATATACATCTCTCAAT 962
DB 1848 TGCTTATTAACATATGCTGAGAAATGACCTTAATTTGATTTGATTTGCTGCACTTT 1907
QY 963 TTCTTATATGATTAATAAAGATATAGAGATTCAATAGAGAAATAGAAATGAAGCAT 1022
DB 1908 TCCTTTTATGATATAGTGTCTTTTCCAGAGAGTGAATTAACATTAACAAGAGGT 1967
QY 1023 TAAGA 1027
DB 1968 TTATA 1972

RESULT 6

US-10-342-821-3
Sequence 3, Application US/10342821
Publication No. US20030154510A1
GENERAL INFORMATION:
APPLICANT: Jana, Wojciechowska
APPLICANT: Evgeny, Lewitin
APPLICANT: Ludmila, Revina
APPLICANT: Igor, Zalunin
APPLICANT: Galina, Cheslukhina
TITLE OF INVENTION: NOVEL DELTA-ENDOTOXINS AND NUCLEIC ACID SEQUENCES CODING THEREFOR
FILE REFERENCE: S-30913C
CURRENT APPLICATION NUMBER: US/10/342,821
CURRENT FILING DATE: 2003-01-15
PRIOR APPLICATION NUMBER: US 60/175,158
PRIOR FILING DATE: 2000-01-07
NUMBER OF SEQ ID NOS: 5
SOFTWARE: PatentIn version 3.0
SEQ ID NO 3
LENGTH: 4896
TYPE: DNA
ORGANISM: *Bacillus thuringiensis*
FEATURE:
NAME/KEY: source
LOCATION: (1)..(4896)
OTHER INFORMATION: product: Cry28a1 delta-endotoxin
PUBLICATION INFORMATION:
AUTHORS: Wojciechowska, et al.
TITLE: Two novel delta-endotoxin gene families cry26 and cry28 from *Bacillus*
TITLE: *thuringiensis* ssp. *finitimus*
JOURNAL: *FEBS Lett.*
VOLUME: 453
ISSUE: 12
PAGES: 46-48
DATE: 1999-06-18
US-10-342-821-3

QY 663 ATTACTACTATTTATGCGAAGCTGCTAATTTTCAATTTAATTTATTAACAAGGTGC 722
DB 1617 ATTATTAAGTTGCTATACAGAGCTGCAGAAATTTACATTTGAAATTTATTAACATCAAGGTGT 1676
QY 723 TGAATGGCTGATGAATGGAATGCAATATACATCCCTCCCAATTAACCTAATGCTGG 782
DB 1677 ACAATTCGCGATCAATGAATGCAATCAACCATTTCCCAAT-----GTTGAA 1727
QY 783 AACATGATGACTATTTAATACTTTTAAAGAAATATATCTTAATATAGTACTATTG 842
DB 1728 GTGATGAGTACTATTTATGACGATTTGATTTATTTGAAAGATATTTATTTTGTG 1787
QY 843 TGCATAATCTTATGAACAGAGCTAAAAATCTTTAGAGCAACCAATATGAATGAG 902
DB 1788 CACCAAGACATACATTAAGGATTTGATTCACCTTAAAGAAATCAAGAAAAATCAATGGA 1847
QY 903 TATATTAATGATGAGATATATATGACATTTCTGTATATACATCTCTCAAT 962
DB 1848 TGCTTATTAACATATGCTGAGAAATGACCTTAATTTGATTTGATTTGCTGCACTTT 1907
QY 963 TTCTTATATGATTAATAAAGATATAGAGATTCAATAGAGAAATAGAAATGAAGCAT 1022
DB 1908 TCCTTTTATGATATAGTGTCTTTTCCAGAGAGTGAATTAACATTAACAAGAGGT 1967
QY 1023 TAAGA 1027
DB 1968 TTATA 1972
Query Match 7.4%; Score 163.8; DB 6; Length 4896;

Best Local Similarity 55.9%; Pred. No. 1.7e-20;
Matches 405; Conservative 0; Mismatches 272; Indels 48; Gaps 3;

```
OY 306 AGGTGTAATTAATATCTTTGGTACCTTAATCACTGCTTTGGGCGGCGGAGAA 365
Db 1293 AGGTGTAATTAATATCTTTGGTACCTTAATCACTGCTTTGGGCGGCGGAGAA 1352
OY 366 AGA-----CAAAAGATAGGACAAATTTAATAATGGAGAAATTTTGGATACCC 422
Db 1353 AGATCCAAAAAATTTGGTCACAATTTATGAACGCGAGAAAGCCTTTAATCAAC 1412
OY 423 GTTAAACAGAAACATPAAACAGCTAAAGTTACAACTTTAGAGATTAGCAAAATAT 482
Db 1413 AATTTCTACAGCTGTAAGAAATGATTAAGTACCTCACTTAAATGTTTAAAGATATAT 1472
OY 483 ACAAGCTAATTAATACGATTAATGATGAGAAATTAATAAGCTCAAGCTCCCTGG 542
Db 1473 AAGCTACTATGAAGAGCATTTAATGATGAGAAATTCAGAGCAAAATATGCGCAG 1532
OY 543 ATTACCAACCATCATGAGCATTAACAAACAGCTGCTGACTTAAATACGATTTGAGAA 602
Db 1533 ATTGGATCAGAG-----GATTTGAAAA 1556
OY 603 TGTTCACATGATTTTATTCGAGAAATACCTGTTTCCAACTTGAAATTTAATAACGCT 662
Db 1557 CGCTCATTTCAATTTTGTAAAGCAATATGCGCAACCTCCCACTGATGACATTT 1616
OY 663 ATTACACCTATTTATGCGCAAGCTGCTAATTTTCACTTAAATTTATTAACAAGGTC 722
Db 1617 AATTAATTAAGTCTATACAGAGCTGCAAAATTTACATTTGAATTTATTAATCAATGAT 1676
OY 723 TGAATTTGCTGATGATGAGATGAGATATACATCTTCAAAATTTGAATGATGCTGG 782
Db 1677 ACAATTCGCGGATGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGAT 1727
OY 783 AACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 842
Db 1728 GTCATCAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1787
OY 843 TGCAATACCTTAATTAACAGGATGATGATGATGATGATGATGATGATGATGATGAT 902
Db 1788 CACCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1847
OY 903 TATATTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 962
Db 1848 TGCTTATTAACATATGCTGCGAAATGATGATGATGATGATGATGATGATGATGATGAT 1907
OY 963 TTTCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1022
Db 1908 TCTTTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1967
OY 1023 TAAAG 1027
Db 1968 TTATA 1972
```

RESULT 7

US-10-089-678-2
; Sequence 2, Application US/10089678
; Publication No. US20030017967A1
; GENERAL INFORMATION:
; APPLICANT: ASANO, Shinichi
; TITLE OF INVENTION: PROTEIN HAVING INSECTICIDAL ACTIVITY, DNA ENCODING THE PROTEIN, A
; TITLE OF INVENTION: NOXIOUS ORGANISM-CONTROLLING AGENT AND METHOD
; FILE REFERENCE: Q68821
; CURRENT APPLICATION NUMBER: US/10/089,678
; CURRENT FILING DATE: 2002-05-02
; PRIOR APPLICATION NUMBER: JP 2000-236140
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: PCT/JP01/06660
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 2
; LENGTH: 3504
; TYPE: DNA
; ORGANISM: Bacillus thuringiensis
; FEATURE:
; NAME/KEY: exon
; LOCATION: (1)..(3501)
; OTHER INFORMATION:
US-10-089-678-2

Query Match 4.5%; Score 98.4; DB 5; Length 3504;
Best Local Similarity 43.4%; Pred. No. 3.1e-08;
Matches 602; Conservative 0; Mismatches 751; Indels 33; Gaps 2;

```
OY 612 TGATTTTATGAGAAATACCTGTTTCCAACTTGAACCTTAATAAGCTTTATAC 671
Db 531 TGATTTTATGAGAAATACCTGTTTCCAACTTGAACCTTGAACCTTAATAAGCTTTATAC 590
OY 672 TATTTATGCGCAAGCTGCTAATTTTCAATTAATTTATTAACAAGGTGCTGAATTTGCG 731
Db 591 AGGTATGACAAAGCAAGCAATTTACATTTGCTATTTAGAGAGATGCTTCCATTTTGG 650
OY 732 TGATGAATGGAATGCAATATACATCTTCAAAATGAACTTAATGCTGAACATCAGA 791
Db 651 AGCAGAGTGG-----GGATTCAGACCGAGGAAATTTTC 683
OY 792 TGACTATTAATACTTTTAAAGAAATATACCTAATTAATGATATGTAATTTGCAAAATAC 851
Db 684 CACATTTTATGATCTCAGGTGACAGTACCGCCCAATATCTGGAATTTATGTTGTAAGTG 743
OY 852 CTATGAAACAGGACTTAAATTTCTTAGAGACGAAACCAATATGAATGAGATATTTAA 911
Db 744 GATATACACTGGCTTAGATTAATTAAGTAAGTAAGTAAGTAAGTAAGTAAGTAAGTAAG 803
OY 912 TGACTATGAAAGATATATGACATTTACTGTTATTAATACATCTTCAATTTTCTTATA 971
Db 804 CCAATTCGAAAGAGAAATGACATTTACTGTTATTAATTAAGTAAGTAAGTAAGTAAGTAAG 863
OY 972 TGATTTTAAAGATATATGAGATTTCAATGAGGAATGAGTAAGTAAGTAAGTAAGTAAGTAAG 1031
Db 864 TGACACAGTACGATTCATGAAATGAAACCAAGCCCACTTACACGGGAAATGATATACAGA 923
OY 1032 ACTCAAGAGAAATTTTATACACTGAAATTAATTTTATGATGCTTCTCTCACTTAGAGT 1091
Db 924 TCCATATATATTTAACAAGAAACAGATGATGATTTTGTAGGCTTGCTGATACCTTAACAG 983
OY 1092 TCAACCAATCTAGCTAGATGAGATTAATTAATTAACAGTGAATTTTAAATTTTTC 1151
Db 984 TGATTTTCTTTTCAAGAGTGAAGCGCTGATTTGTTTCAACACCTATTTGATAT 1043
OY 1152 ATTTTTAAACAATTTATTTTATTAAGAAATTAATTAATTTTGGGAATTCGTTAGTTGG 1211
Db 1044 ACTCAGTAAATTAAGATTTTATTAACAAGAGCGGGCTTCCCTTGAATTAATGGAATA 1103
OY 1212 TATTTCTAATGCTGATGACCTACT-----TATAGCAATATCTAATATGAACTTTATA 1265
Db 1104 CTTGAATATTTGGTAGAGACATTTCTAATAATTAATAATTAATGATGCTGATACGANTT 1163
OY 1266 TGGAGAAAGAAAGGTTCAACCAACAACAAATTAAGACATTTGAATCTTAAAGT 1325
Db 1164 AGAAGCTAATTAACGCTAGATTAATCTTAACAAATCAAGTATTAATTAAGCAAAATTA 1223
OY 1326 TTCAATGTAATGATTAACAATCACTCTGTTTCCCTTAATGCAACCACTTTATATAT 1385
Db 1224 GGAATCTTTCAAGGTCATCATTTAGGGCGGATTTAGCTAATTAATTAATTAATTAAT 1283
OY 1386 TAATCAATTAATTAATTTAATTTGCTCACTTAACAACACATCAATATTTAGCAGAG 1445
Db 1284 TGAAGTTCGATAGCTATTTTATACAGCTTGAACAAGATACAGATCAGATGCTGG 1343
OY 1446 AGGCTTTATCTAATTAATTAACAACAACATTTTCTTAATTTCTTAAGAAAAAAGACTG 1505
Db 1344 AGTTTATGTAATCAAAACCAATACATTAATGAGATGATGATGATGATGATGATGATGAT 1403
```


OY 1506 CAATCTAGTATGATCCAGGTTGTTCCAAACTTATATACATATAGTATATTTATC 1565
DB 1404 GATTGATGAATCCCTCCAGAGAAATGAGCCACTTAGTAGAGGATATACCATATATATC 1463
OY 1566 CCATTTTTCATTTATTTACTTATTTCTATATGATGATTTAGACCTACCAATATTAAGATAC 1625
DB 1464 TCATATACCTCTTATTTCTTTTCTTAAAGATGATAGTCTCTGCTAGATATGCAATCT 1523
OY 1626 AGGTATATAGATGAGACACAGATGATGATATATATATGCAATATGCAATTAAT 1685
DB 1524 CCGTATATTTGCTTGGACACATCGAGTGGGATGTTCAAAATACATTTATTCAGATTA 1583
OY 1686 AATTACAAATGATCCAGCAATCAAGATTAACATCTTGAATCAAACTTAAAGTATGA 1745
DB 1584 AATTACTCAGATACCAAGTTGTAAGGACATATCTTAGTTTCAAGTACTACTGTTATTA 1643
OY 1746 AGGACCTGCTATACAGAGAACTTGTATTTATTAACAAGGCGCTTAAAGAT 1805
DB 1644 AGGTCTGATTTTACAGAGCAATATCTTAAAGAAAGATGATGCTGTTAGCTTA 1703
OY 1806 TACATGTAACTCCTAATTTCTACCAATCTTATTTCACTAGACTTGTGATATGCTACAA 1865
DB 1704 TACTAGTCTCTGTAATATCACTTATTCACAAAGATATGTCGACAGAAATACGTTATGC 1763
OY 1866 TGGTGTGGAATATCTTCTTAATATATCTTTACATACAGAGATTAAGAAATCC 1925
DB 1764 TTCTACTACTAATCTTACAGACTTTTGTAAATTTCTGAACTGCAATTTACTATTA 1823
OY 1926 ACCTCACAGATCAACAACATTTTCTGTGACAAATTAATTAATTAACAATACGAGA 1985
DB 1824 TGTTAATTAACCAATGAATTAAGGAGATGATTAACATTTAATGATTAAGTAAAC 1883
OY 1986 TTTTGG 1991
DB 1884 TATTGG 1889

RESULT 8
US-10-089-678-3
; Sequence 3, Application US/10089678
; Publication No. US20030017967A1
; GENERAL INFORMATION:
; APPLICANT: ASANO, Shinichiro
; TITLE OF INVENTION: PROTEIN HAVING INSECTICIDAL ACTIVITY, DNA ENCODING THE PROTEIN, A
; FILE REFERENCE: 068821
; CURRENT APPLICATION NUMBER: US/10/089,678
; PRIOR FILING DATE: 2002-05-02
; PRIOR APPLICATION NUMBER: JP 2000-236140
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: PCT/JP01/06660
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 3
; LENGTH: 3690
; TYPE: DNA
; ORGANISM: *Bacillus thuringiensis*
US-10-089-678-3

Query Match 4.5%; Score 98.4; DB 5; Length 3690;
Best Local Similarity 43.4%; Pred.No. 3.1e-08;
Matches 602; Conservative 0; Mismatches 751; Indels 33; Gaps 2;

OY 612 TGATTTTATTTGAGAAATACCTGTTTCCAACTGAAACTTATTAACGCTATTAATCC 671
DB 717 TGATTTTGTCTATAAATCCATCTTTTGCATATATCGACAGAAAGTACATTAATATC 776
OY 672 TATTATGCGAGAGTGTCTAATTTTCAATTAATTAATTAACAAGGTCGTAATGGC 731
DB 777 AGTGTATGACAAAGCAAGAAATTTACATTTGCTATTTATGAGATGCTTCCATTTTGG 836

OY 732 TGATGAATGGAATGACAGATATACATCTTCACAAATTTGAACCTAATGCTGGAACATGAGA 791
DB 837 AGCAGAGTGG-----GGAATTCACACAGAGAAATTTTC 869
OY 792 TGACTATTTATTAATCTTTTAAAGAAATATACCTAATATATATATATATATATATATATAT 851
DB 870 CACATTTTATGATGTCAGAGTGACAGTACCGCCCAATACCTGGAATTAATGATTAAGTGG 929
OY 852 CTATAGAACAGAGATTAATAAATCTTAGAGACGAACCAATATGAAATGAGATATATTTAA 911
DB 930 GTATTAACCTGCTTATATTAATTAATAAGATACGAATGCTGCAAGTTGGCTGAAGATATCA 989
OY 912 TGACTATGGAATATATGACATTAATGATTAATGATTAATGATTAATGATTAATTTCTTTATA 971
DB 990 CCAATTCGGAAGAAATGACATTAATGATTAATGATTAATGATTAATGATTAATTTCCAACTA 1049
OY 972 TGATTTTAAAGATATGAGTTTCAATATGAGAGAAATGAAAGTAAAGGATTAAGATGA 1031
DB 1050 TGACACAGATGATGATCAATGGAACACCGCCCACTTACACGGAAGGTATATACGA 1109
OY 1032 ACTCACAAGAAATTTATACAACTGAAATTAATTTGATGCTGCTCTCACTTAGAGT 1091
DB 1110 TCCATATGATATTTAACAAGAAACAGTGTGATTTTGTAGGGTGTGCTACTTAACAG 1169
OY 1092 TCAACCCAACTTACGATGATGATTAATTTAACAAGTGAAGTTTAATTAATTTATTC 1151
DB 1170 TGATATTTCTTTTACAGAGTGAAGCGCTGTATTCGTTTCAACACACCTATTTGATAT 1229
OY 1152 ATTTTGAACAAATTTATTTTATACGAATAATGCAATTTTGGGAATGCTTTAGTTGG 1211
DB 1230 ACTCAGTGAATATGAATTTTATATACAAAGACGGGCTTCCCTTGAATATATACGAATA 1289
OY 1212 TATTTCAATCGTATGACACTACT-----TATAGCAATCTTAACTGAACTTTATA 1265
DB 1290 CTTGAATATTTGGTATGAGACTTTCTATTAATATATTAATATGCAATGCTCATGACATT 1349
OY 1266 TGAAGAAAGAACAGTTTCAACCCACACAAACAAATTAAGCAATTTGAATCTTAAT 1325
DB 1350 AGAAGTAAATTAAGCGATGATTAATCTTAACAAATCAAGATTAATTAATTAAGCAATA 1409
OY 1326 TTCAATTTGATGATGACATGACATGCTGTTTCCCTATTTCAACACACTTATTAAT 1385
DB 1410 GATATATCTTCAAGTTGATGATCAATTAAGGGGCGATTAAGCTAATTAATCAACGATTA 1469
OY 1386 TAATCAATTTGAATTTATTTTAAATGCTCATCAACAACACTCAATTAATTCAGCAGG 1445
DB 1470 TGAAGTCCGATGACCTGTTTACATGCTTGAACAAGATTAAGATCAAGATCAAGTTGG 1529
OY 1446 AGGTCTTATCTAATTAATCAAAACACAACTTTTTCATTTCTAGAAAAAAGACTG 1505
DB 1530 AGGTTTACGTACTCAAAACCAATACACTATGCAAGTATGTCACAAATTTCAATAC 1589
OY 1506 CAATCTAGTATGATCCAGGTTGTTCAACAACTTAAATTAATCAATGATATATTTATC 1565
DB 1590 GATGATGAATATCCCTCAGAGATGAGCCACTTAGTAGAGGATGATGATCAATATATC 1649
OY 1566 CCATTTTCAATTTATTTACTTATTTCTATGATTTGATTAACAGCTACAAATATTTAGATAC 1625
DB 1550 TCATATACCTCTTATTTCTTTTCTAAGAAATGCTAGTAGTCTGCTAGATATGCAATCT 1709
OY 1626 AGGTATATAGATGAGACACAGATGATGATATATATATGCAATATATCAATTAAT 1685
DB 1710 CCGTATTTGCTTGGACACATCGAGTGGGATGTTACAAATATCAAGTTTATTCAGATTA 1769
OY 1686 AATTACATGATCCACCAATCAAGATTAACATCTTGAATTAACAACCTTAAGATTAATGA 1745
DB 1770 AATTACTCAGATACCAATTTGTAAGGACATATCTTATGTTTCAAGTACTCTGTTATTA 1829
OY 1746 AGGACCTGCTCATACAGAGAACTTGTATTTTCAAAAGTCAAGGCGTATTAAGAAAT 1805
DB 1830 AGGTCTGGAATTAAGAGGCAATATCTTTAAAGAAACAGATGATGCTGCTTATGCTTA 1889
OY 1806 TACATGTGAATCTCTAATTTCTACCAATCTTATTTCAATTAAGCTTGAATATGCTACAA 1865


```
Db 1890 TACTAGTCTCTGTAAATACACATTTATCAAAAGATATGCAAGAAATCGTTATGC 1949
Qy 1866 TGGTCTGGAATATCTCTCCAAATATATCTCTCAATACAGAGTAATGAAATCC 1925
Db 1950 TTCTACTACTACTTACGACCTTTTGTGAACAATTTCTGCACTCCATTTACTCTATTA 2009
Qy 1926 ACTCAAGCACTCAACAACTTTTCTGTACAAATTAATTAATTTTACATACGAGA 1985
Db 2010 TGTATATTAACCAATGAATTAAGGGAGATTTAACATTATACATTGACTTAGCAAC 2069
Qy 1986 TTTTGG 1991
Db 2070 TATTGG 2075

RESULT 9
US-10-312-841-1/c
; Sequence 1, Application US/10312841
; Publication No. US20030186277A1
; GENERAL INFORMATION:
; APPLICANT: Bpigenomics AG
; TITLE OF INVENTION: Diagnose von bedeurenden genetischen Parametern innerhalb de s MHC
; FILE REFERENCE: E01/1208/WO
; CURRENT APPLICATION NUMBER: US/10/312,841
; NUMBER OF SEQ ID NOS: 2
; SEQ ID NO 1
; LENGTH: 3673778
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
; NAME/KEY: unsure
; LOCATION: (3294164)
; US-10-312-841-1

Query Match 4.4%; Score 97; DB 6; Length 3673778;
Best Local Similarity 43.1%; Pred. No. 1.1e-06;
Matches 629; Conservative 0; Mismatches 825; Indels 7; Gaps 3;

Qy 749 ATATACATCTTCACAAATGGAACCTTAATGCTGGAACATCAGATGACTATTATAACTTT 808
Db 1715403 AAATATATATATTAATAATACATTAATAATATATATATTAATAATATACATTAATAATAT 1715344
Qy 809 TAAAGAAATATATCTTAATATAGTAATCTATGTGCAATTCCTATAGAAACGAGACTPA 868
Db 1715343 ATATATATTAATAATATATATTAATAATATATATATTAATAATATATATATATAT 1715284
Qy 869 AAATCTTAGAGAGCAACCAATATGAATGAGATATTTTAATGACTATGGAAGATTA 928
Db 1715283 AAAATATATATATTAATAATATATATTAATAATATATATTAATAATATATTAATAATTA 1715224
Qy 929 TGACCATTTACTGTATTTAGATACCATCTCTCAATTTTCTTTATATGATTTAAAGATATA 988
Db 1715223 TATATTAATAATATATATATATTAATAATATATATTAATAATATATATATATATATAT 1715164
Qy 989 GAGATTCATAGAGGAATAGAAAGTAAAGGATTAAGATGAACCTCAACAGAGAAATTT 1048
Db 1715163 TAAATATATATATTAATAATATATATTAATAATATATATTAATAATATAT 1715104
Qy 1049 ATACAACTGAATTAATTTTGAATGCTCTTCCTCAACTTGAAGCTCAACCAATCTAGCTA 1108
Db 1715103 ATATATATTAATAATATATATATTAATAATATATATTAATAATATATATATTAATAAT 1715044
Qy 1109 CGATGAATATATATTTAAACAGTCGCAAGTTTAAATTTTTCATTTTAAACAATTTA 1168
Db 1715043 ATATATATTAATAATATATATATTAATAATATATATTAATAATATATATATAT 1714984
Qy 1169 TTTTATATACAGAAATACAAATTTGCGAATCGTTAGTGGTATTTCTAATCGTGATG 1228
Db 1714983 ATATATATATTAATAATATATATATATATATATATATATATATATATATATATAT 1714927
```

```
Qy 1229 CACCTACTTAATAGCAATCTATTAACCTGAACCTTTATATAGAGAAAGAACAGGTCACCCA 1288
Db 1714926 TATATATATTAATAATATATATTAATAATATATATATATATATATATATATATATAT 1714867
Qy 1289 CAACAAAACAAATAAGACCAATTTGAATCTTATA--AGTTCAATTTGTAAGTATAGC 1345
Db 1714866 AAATATATACATATTAATAATATATATATATATATATATATATATATATATATATAT 1714807
Qy 1346 AATGACCTCCGTTCCCTTCCCTTTCACACCACTTTATATATTAATCAAAATGCACTTTAT 1405
Db 1714806 AATATATATTAATAATATATATATTAATAATATATATATATATATATATATATAT 1714747
Qy 1406 TAAATGCTCATCTTAACAACACACATCAATATTTACGAGAGGGCTTTATCTAATTTATC 1465
Db 1714746 TAAATATACATATTAATAATATATATATTAATAATATATATATATATATATATATAT 1714687
Qy 1466 AAAACACAACTTTTTCATTTCTGAGAAAAAAGACTGCAATCTAGTTATGATCCAG 1525
Db 1714686 ACATATTAATAATATATATATATATATATATATATATATATATATATATATATAT 1714627
Qy 1526 GTTGTCACCAACTTTAATACTATAGTCATATTTTATCCATTTTCATATTTACTT 1585
Db 1714626 AATATATATTAATAATATATATATATATATATATATATATATATATATATATAT 1714567
Qy 1586 ATTCCTATGTGATTTGATTTACAGCTACCAATATTTAGATACAGGTGATATAGATGACAC 1645
Db 1714566 TATACATATTAATAATATATATATATATATATATATATATATATATATATATAT 1714507
Qy 1646 ACAGTATGTGATAGATATATATATATATATATATATATATATATATATATATATAT 1705
Db 1714506 ATATATATTAATAATATATATATATATATATATATATATATATATATATATATAT 1714447
Qy 1706 TCAAAAGTAACTTTGATATCAAACTTAAGTATTTGAAGAGCTGCTGATACAGAG 1765
Db 1714446 TATATATTAATAATATATATATATATATATATATATATATATATATATATATAT 1714387
Qy 1766 GAACTTGTTTATTTTCAAAAGTCAAGGGCGTTTGAATTAATCAATGTGAACTCTTAAT 1825
Db 1714386 ATATATTAATAATATATATATATATATATATATATATATATATATATATATAT 1714327
Qy 1826 CTACCAATCTTAATTTCTATGACTTGCATATGCTACAAATGTGCTGGAATATCTTTC 1885
Db 1714326 TATATATTAATAATATATATATATATATATATATATATATATATATATATAT 1714267
Qy 1886 CTATATATCTCTTACCAATACAGAGTAATATAGAAATCCACTCAACAGCTCAACACA 1945
Db 1714266 ATTAATAATATATATATATATATATATATATATATATATATATATATATAT 1714207
Qy 1946 CTTTTCTGATCAAAATATATATATATATATATATATATATATATATATATATATATAT 2005
Db 1714206 ATATATTAATAATATATATATATATATATATATATATATATATATATATATAT 1714148
Qy 2006 CAACTGACATTAACCTTTAAATGCAACATACCACTTATATTTTAATCGTCAGATG 2065
Db 1714147 ATATTAATAATATATATATATATATATATATATATATATATATATATATATAT 1714088
Qy 2066 TATCAATTTCAATTTTATCTATGATTAATAATTTGAATTTATCCAAATTTCTCCTATAGC 2125
Db 1714087 TATTAATAATATTAATAATATATATATATATATATATATATATATATATATATATAT 1714028
Qy 2126 ACCAAATATAGGAGAAAAAATAAATTTAGAACTATCCAAACAAATAAATATATATTTCA 2185
Db 1714027 ATTAATAATATATATATATATATATATATATATATATATATATATATATATATAT 1713968
Qy 2186 CAAATCATACAAAAACACTTT 2206
Db 1713967 TAAATATATATATATATATATATATATATATATATATATATATATATATATAT 1713947

RESULT 10
US-10-311-455-2128/c
; Sequence 2128, Application US/10311455
```



```
; APPLICANT: Payne, Jewel
; APPLICANT: Sick, August J.
; TITLE OF INVENTION: Novel Bacillus thuringiensis Isolate Active Against Lepidopteran
; TITLE OF INVENTION: Pests, and Genes Encoding Novel Lepidopteran-Active Toxins
; FILE REFERENCE: MA-43CDF204
; CURRENT APPLICATION NUMBER: US/10/825,751
; CURRENT FILING DATE: 2004-04-16
; PRIOR APPLICATION NUMBER: US 09/837,961
; PRIOR FILING DATE: 2001-04-19
; PRIOR APPLICATION NUMBER: US 09/521,344
; PRIOR FILING DATE: 2000-03-09
; PRIOR APPLICATION NUMBER: US 08/933,891
; PRIOR FILING DATE: 1997-09-19
; PRIOR APPLICATION NUMBER: US 08/356,034
; PRIOR FILING DATE: 1994-12-14
; PRIOR APPLICATION NUMBER: US 08/210,110
; PRIOR FILING DATE: 1994-03-17
; PRIOR APPLICATION NUMBER: US 07/865,168
; PRIOR FILING DATE: 1992-04-09
; PRIOR APPLICATION NUMBER: US 07/451,261
; PRIOR FILING DATE: 1989-12-14
; PRIOR APPLICATION NUMBER: US 371,955
; PRIOR FILING DATE: 1989-06-27
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 7
; LENGTH: 3522
; TYPE: DNA
; ORGANISM: Bacillus thuringiensis
; US-10-825-751-7
```

Query Match 3.9%; Score 85.8; DB 8; Length 3522;

Best Local Similarity 53.8%; Pred. No. 7.3e-06;

Matches 220; Conservative 0; Mismatches 162; Indels 27; Gaps 1;

```
QY 589 ATAGATTGGAGATGTTTCAAGATTTTATTCGAAATACCTGGTTCCACTTGA 648
DB 388 ATTCGATTGGCTAATACAGACGCTTAAATACAGAAATTAATTTACACTTACA 447
QY 649 ACTTATTAAGAGCTTATCTACTATTTATGGCGAAGCTGCTAATTTTCAATTAATT 708
DB 448 AGTTTGAATATCCCTCTTTTATCGGCTATGTTCAAGGCGCAATTTTCAATTTACACT 507
QY 709 TTACAACAGAGTGTGAATGGCTGATGAATGAGATATACATCTTCAACAAT 768
DB 508 TTAAGAGACGCTGATTCGTTGGGAGGCTTGGGACGATATAGCTACTTAAAT 567
QY 769 GAACCTAATGCTGAACATCAGATGACTATTAATACTTTTAAAGAAATATACCTAAA 828
DB 568 C-----ATTATATAGATTAATAATCTTATTCATAGA 600
QY 829 TATAGTACTATTTGCGAAATACCTATAGAACAGACTTAAATCTTATAGAGACAGACA 888
DB 601 TATAGGAACAATGTTTGGACACATACATCAAGATTAAGAACTTAAGAGTACTAAT 660
QY 889 AATATGAATGAGATATTTAATGACTATCGAAGATATATGACCACTTACTGTATTAGAT 948
DB 661 ACTGACAAATGGGCAAGATTCATCATGTTTAGAAGATTTTACCTTACTGTATTAGAT 720
QY 949 ACCATCTCTCAATTTCTTTATATGATATAAAAGATATAGAGATTCA 997
DB 721 ATCGTGTCTCTTTTCCGAACTAGATGTTAGAACATATCCAAATTCAAA 769
```

RESULT 15

US-10-311-455-1280/c

; Sequence 1280, Application US/10311455

; Publication No. US20030143606A1

; GENERAL INFORMATION:

; APPLICANT: ODEK, Alexander

; APPLICANT: PIERENBROCK, Christian

; APPLICANT: BERLIN, Kurt

; TITLE OF INVENTION: Diagnosis of Diseases Associated with the Immune System by Determination of

```
; TITLE OF INVENTION: cytosine methylation
; FILE REFERENCE: 5013.1014
; CURRENT APPLICATION NUMBER: US/10/311,455
; CURRENT FILING DATE: 2002-12-16
; PRIOR APPLICATION NUMBER: PCT/EP01/07537
; PRIOR FILING DATE: 2001-07-02
; PRIOR APPLICATION NUMBER: DE 10032529.7
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: DE 10043826.1
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 2424
; SEQ ID NO 1280
; LENGTH: 6175
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
; US-10-311-455-1280
```

Query Match 3.7%; Score 81.2; DB 6; Length 6175;

Best Local Similarity 42.7%; Pred. No. 6.8e-05;

Matches 593; Conservative 0; Mismatches 788; Indels 9; Gaps 3;

```
QY 800 ATAACTTTTAAAGAAATATACCTAAATATAGTAACTATTTGCAAAATCTATAGAA 859
DB 5961 AAATTAATATCTAATAATATATATATAAACAAGTATTAACATACAAATATAATACAT 5902
QY 860 CAGGACTTAAATCTTGAACAGACCAATATGAAATGAGATATTTATGACTATG 919
DB 5901 AAACAAATATATATATACATACGTAACAAATATTAACAAATATTAATATATATAT 5842
QY 920 GAAGTATATGACCATTTCTGATTTAGATACATCTCTCAATTTTCTTATATGATATTA 979
DB 5841 ATAAATTAACAAATATATCAATTAATTAATTAATTAATTAATTAATTAATTAAT 5782
QY 980 AAAGTATATGAGATTCATTAAGAGAAATAGAGTAAAGCAATTAAGATGAATCAGAA 1039
DB 5781 AAACACTATTAATATATATATATATATATATATATATATATATATATATATATAT 5722
QY 1040 GAGAAATTTATACACATGAATTAATTTGATCGCTCCCTCACTTGAAGTCAACCA 1099
DB 5721 ATAAACATATATATATATATATATATATATATATATATATATATATATATATAT 5662
QY 1100 ATTAGTACGATGAATATATATATATATATATATATATATATATATATATATATAT 1159
DB 5661 AACATATATATATATATATATATATATATATATATATATATATATATATATATAT 5602
QY 1160 AACATTTATTTTATATACAGAAATATCAAAATTTGCGAATCGTTAGTGTATTTCTA 1219
DB 5601 AAATATATTAATATATATATATATATATATATATATATATATATATATATATAT 5542
QY 1220 ATCGTATGACCTTATATATATATATATATATATATATATATATATATATATATAT 1279
DB 5541 ACATTAATATATATATATATATATATATATATATATATATATATATATATATAT 5482
QY 1280 GTTCAACCAACAAACAAATATAGACATTTGAATCTTATATATATATATATATATAT 1339
DB 5481 AACATATATATATATATATATATATATATATATATATATATATATATATATATAT 5422
QY 1340 ATAGCAATCACTCTGTTTCCCTATTCACCAACACTTATATATATATATATATATAT 1399
DB 5421 AAAATATATATATATATATATATATATATATATATATATATATATATATATATAT 5362
QY 1400 TTTATTTAATGGCTCATCTAACAACACTCAAAATATTCAGAGAGAGGCTTATATCTA 1459
DB 5361 TATATATATATATATATATATATATATATATATATATATATATATATATATATAT 5306
QY 1460 ATTAATCAAAACAACTTTTTCATTTCTAGAAAAAAGACTGCATCTAGTTATTTG 1519
DB 5305 AATATATATCGTACAAATATATATATATATATATATATATATATATATATATATAT 5246
QY 1520 ATCCAGGTTGTCACCAACTTATATATATATATATATATATATATATATATATATATAT 1579
```


GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 10, 2005, 18:16:09 ; Search time 397 Seconds
(without alignments)
9886.287 Million cell updates/sec

Title: US-10-783-417-1

Perfect score: 2208

Sequence: 1 atgatacaataacagataa.....atcatacaaaacacttga 2208

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 1303057 seqs, 888780828 residues

Total number of hits satisfying chosen parameters: 2606114

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents NA:*

- 1: /cgn2_6/ptodata/1/ina/1 COMB.seq:*
- 2: /cgn2_6/ptodata/1/ina/5 COMB.seq:*
- 3: /cgn2_6/ptodata/1/ina/6A COMB.seq:*
- 4: /cgn2_6/ptodata/1/ina/6B COMB.seq:*
- 5: /cgn2_6/ptodata/1/ina/H COMB.seq:*
- 6: /cgn2_6/ptodata/1/ina/PCTUS COMB.seq:*
- 7: /cgn2_6/ptodata/1/ina/PP COMB.seq:*
- 8: /cgn2_6/ptodata/1/ina/RE COMB.seq:*
- 9: /cgn2_6/ptodata/1/ina/backfile1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	340.6	15.4	3543	3	US-09-224-024-27
2	340.6	15.4	3543	6	PCT-US94-07902-27
3	188.2	8.5	2061	3	US-09-224-024-30
4	188.2	8.5	2061	6	PCT-US94-07902-30
5	98.4	4.5	3504	3	US-10-089-678-2
6	98.4	4.5	3504	3	US-10-089-678-3
7	85.8	3.9	3444	2	US-08-349-867-22
8	85.8	3.9	3444	2	US-08-349-867-25
9	85.8	3.9	3444	2	US-08-239-476-22
10	85.8	3.9	3444	2	US-08-239-476-26
11	85.8	3.9	3444	2	US-08-598-305A-22
12	85.8	3.9	3444	2	US-08-598-305A-26
13	85.8	3.9	3444	2	US-08-639-923A-22
14	85.8	3.9	3444	2	US-08-639-923A-26
15	85.8	3.9	3444	6	PCT-US95-05431-22
16	85.8	3.9	3444	6	PCT-US95-05431-26
17	85.8	3.9	3450	2	US-08-349-867-20
18	85.8	3.9	3450	2	US-08-239-476-20
19	85.8	3.9	3450	2	US-08-598-305A-20
20	85.8	3.9	3450	6	PCT-US95-05431-20
21	85.8	3.9	3450	6	PCT-US95-05431-20
22	85.8	3.9	3522	2	US-07-828-788A-9
23	85.8	3.9	3522	2	US-08-349-867-24
24	85.8	3.9	3522	2	US-08-349-867-28

25	85.8	3.9	3522	2	US-08-239-476-24	Sequence 24, Appl
26	85.8	3.9	3522	2	US-08-239-476-28	Sequence 28, Appl
27	85.8	3.9	3522	2	US-08-356-034-7	Sequence 7, Appl
28	85.8	3.9	3522	2	US-08-598-305A-24	Sequence 24, Appl
29	85.8	3.9	3522	2	US-08-598-305A-28	Sequence 28, Appl
30	85.8	3.9	3522	2	US-08-639-923A-28	Sequence 28, Appl
31	85.8	3.9	3522	2	US-08-639-923A-28	Sequence 28, Appl
32	85.8	3.9	3522	3	US-09-178-252-5	Sequence 7, Appl
33	85.8	3.9	3522	3	US-09-178-252-5	Sequence 7, Appl
34	85.8	3.9	3522	3	US-09-521-344-7	Sequence 5, Appl
35	85.8	3.9	3522	3	US-09-826-660-5	Sequence 5, Appl
36	85.8	3.9	3522	3	US-09-837-961A-7	Sequence 7, Appl
37	85.8	3.9	3522	6	PCT-US92-11337-9	Sequence 9, Appl
38	85.8	3.9	3522	6	PCT-US95-05431-24	Sequence 24, Appl
39	85.8	3.9	3522	6	PCT-US95-05431-28	Sequence 28, Appl
40	85.8	3.9	3522	9	5188960-7	Patent No. 5188960
41	85.8	3.9	4020	6	PCT-US91-02560-1	Sequence 1, Appl
42	76.4	3.5	3501	2	US-08-448-170-5	Sequence 5, Appl
43	76.4	3.5	3501	3	US-08-961-803-4	Sequence 4, Appl
44	76.4	3.5	3507	3	US-09-661-322A-21	Sequence 21, Appl
45	76.2	3.5	3504	2	US-08-291-368-3	Sequence 3, Appl

ALIGNMENTS

RESULT 1
US-09-224-024-27
Sequence 27, Application US/09224024
Patent No. 6056953
GENERAL INFORMATION:
APPLICANT: Leslie Hickie
TITLE OF INVENTION: Materials and Methods for the Control of
NUMBER OF SEQUENCES: 31
CORRESPONDENCE ADDRESSES:
ADDRESSEE: David R. Saliwanchik
STREET: 2421 N.W. 41st Street, Suite A-1
CITY: Gainesville
STATE: FL
COUNTRY: USA
ZIP: 32606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/224,024
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/856,226
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Saliwanchik, David R.
REGISTRATION NUMBER: 31,794
REFERENCE/DOCKET NUMBER: MA79
TELECOMMUNICATION INFORMATION:
TELEPHONE: 904-375-8100
TELEFAX: 904-372-5800
INFORMATION FOR SEQ ID NO: 27:
SEQUENCE CHARACTERISTICS:
LENGTH: 3543 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-09-224-024-27
Query Match 15.4%; Score 340.6; DB 3; Length 3543;
Best Local Similarity 53.1%; Pred. No. 5.5e-59;

Matches 1170; Conservative 0; Mismatches 934; Indels 99; Gaps 17;
QY 10 AATAAGATAAACAAGATGAAATTAATGATCGATACCTTAATTTTCCGAC 69
Db AATCCCTATCAAAAATGAAATATGAAACATTAATGCTTCACAAAAAATTAAT 63
QY 70 AGAAGCATATGATTCAGATACCCCTTACACAATAATCCAAATCCACTTCCAAAC 129
Db ATATCTAATTAATTAACAAGATATCCAAATGAGTCCAAAACAATTTATCAAGT 123
QY 130 ACAAAATTAACAAGATGCTCAATATGTGTCAAGGATATACAAATATGTGATATTC 189
Db ACAATATTAAGATGGCTCAATATGTGTCAAGGATATCAAGTATGTGTGAGATTT 183
QY 190 GAGACATTTGCTAGTGTGATACAAATGCTGCAAGTATGTGAGATCTATTGATCCGGT 249
Db GAACTTTATATGAT-----AGTGTGAACCTGAGTGCCTTAATTAATGTAGTGG 234
QY 250 ACTCTGTAGCCGGATAGTGGGCTCATCTATATCCGACCGATAGGAATATAGT 309
Db ACCGACTGACTGGT-----TTCGGGTTCAACAACCCCTTAGGA 273
QY 310 GCTAATATATATCTTTTGTGTAACCTATCACTGTCTTTTGCCGCGGAGAACAGAC 369
Db CTGCTTTAATAGTGTGTGTGATATTAACAGTTCCTTTTCCAGCCCAAGACCAATCT 333
QY 370 AAAACAGTATGACAACTTATTAATAATGGAGAAATTTTGTGATACACCGTTACA 429
Db AACAC---ATGAGAGTACTTTATTAACAACATAAATATTAATTAAGAAATAGCA 390
QY 430 -GAAAGCATAAACGCTAAAGTTACAACTTTAGAGGATTTAGACAAATTTTACAAG 488
Db TCAACATATATAGTAATGCTAATTAATAATTTTAAACAGTGTGTTAATGTATACGACT 450
QY 489 CTATTAATCAGCATTAATGATTTGAGAAATTAATAAAGCTACAGCTCTGATTAAC 548
Db TATCATATATACCTTAATAACATGGAGAAATATCCAAACCAACAATAATCTCAGAGTGA 510
QY 549 ACCATCATACGATTAACAACAGCTGCTTGAATCTTAAATTAACGATTTGAGAAATGTTCA 608
Db AGGA-----CACAAATCCAGTAGTTCATTAACCATTTTCAAAATGTCAATTCAGAGCTTGT 566
QY 609 CAATGATTTTATGAGAAATACCTGTGTTCCAACTTGAACCTTAATAAGCTTATTAATCT 668
Db AAACCTGTGTCTCTTAATCTCTAGTATGTC-----GATTAATTAATCAATCACTAGATTT 620
QY 669 ACCTATTTATGCGAAGCTGCTAAATTTTCAATTAATTAATTAACAACAGGTCGTAAT 728
Db ATCTAGTTATGCAACAGCAACAACTTACATCTGATTAATTAATCAAGCCGTCAAAATTT 680
QY 729 GAGTGAATGAAATGCAATATATCATCTTCAAAATTTGAACCTTAATGCTGGAACATC 788
Db TGAAGCGTATTTAAAAACAATGCAATGCAATGATTAATTAAGAGCTTTGCC---AACGCG 737
QY 789 AGATGATCTATTAATCTTTTAAAGAAATATATCTTAATTAATTAATTAATTAATTTGTGCAA 848
Db AATGATTAATTAATCCAGTATTAATGCTTAAGAGATTAATCACTAATTAATTTGTGTAAC 797
QY 849 TACCTATGAGACAGACTTAATAAATCTTAAGAGAGAACTTAATTTGAATGAGATATTT 908
Db AACTTATTAATAAGATTAATTAATTAATTAATAACGAGCTGATAGTATCTTGAATGAGAA 857
QY 909 TAATGATCTATGAAAGATTAATGACATTAATCTGATTAATGATCAATCTCTCAATTTTCTTT 968
Db TATAAATCTGAAACATA-----CAATACGATTCGAACAAATAATGACTACTGCTGATTT 911
QY 969 ATATGATTAATAAAGATTAATGATTAATGATTAATGAGGATTAAGATTAAGATTAAGAA 1028
Db AGATCTGTGTGACTCTTTCTTAATTAATGATGATGATTAATTAATCAATAGGTGTCATATC 971
QY 1029 TGAATCTCAAGAGATTTTATACAACTGAATTAATTTGATGCTTCTCTCAACTTAG 1088
Db TGAATCTCAAGAGATTTTAT--CAGGTACTTAATCTTGAAGAAAGCCCTTAATTAATAT 1029

QY 1089 AGTTCAACCAATCTAGCTACGATGGAATATTAATTAACAAGTGCAGATTTTAATTAAT 1148
Db TATGACTTTCAATATCAAGAGATTCATCTTACAGTGAACC--GCATTTATTAATCTTGGC 1087
QY 1149 TTGATTTTGAACAATTTATTTTATTAAGAGAAATTAACAAATTTGGGAATGCTTAGT 1208
Db TTGATTTTGA-----TTTTATGAAAAAGGCAACCTACTCTAATTAATTTTTTCAAC 1142
QY 1209 TGGTATTTCTAATCGTATGACACCTTAATTAAGCAATTAATTAATGAACTTTATATAG 1268
Db CAGCATTAATTAATATGTTTCAATTAACACTTGAATTAATTAATCCAAAAATCTAGTGT 1202
QY 1269 AGAAGAACAGGTTCAACCAACAACAAATTAAGACATTTGAATCTTAATTAAGTTTC 1328
Db TGGAAATCACAATGATATGATTAATTAATCTTGTGTTGGCAAAATATTTATAT 1262
QY 1329 AATGTTAATGATTAAGCAATCACTCTGTTTCCCATTAATCAACACACTTAATTAATTA 1388
Db TTTTATTAATTAATGATCAATAGCTTAATTAATTAATTAATTAATTAATTAATTAATAG 1322
QY 1389 TCAAAATGAATTTATTTAATGAGCTCATCTTAACAACAACCTCAATATTCAGCAGAGAG 1448
Db TAAATGATTTTATTAATTAATTAATGATTAATGATTAATTAATTAATTAATTAATTAAT 1382
QY 1449 GTCT-----TTATCTAATTAATCAAAACACAACTTTTTCATTTTCTAGAAAAAAGA 1502
Db ATCTGGCAATTAATTAATTAATGATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1442
QY 1503 CTGCAATCTGATTAATTAATGATTAATGATTAATTAATTAATTAATTAATTAATTAAT 1562
Db AAGAGAAATCAAGCAATCCCTTACCTTTTCCAAATTAATTAATTAATTAATTAATTAAT 1502
QY 1563 ATCCCATTTTCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1622
Db ATC-----ATTATTAATAAGCTTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1545
QY 1623 TACAGGTGA--TTAGATGACACACAGTATGTTGATTAATTAATTAATTAATTAATTAAT 1680
Db CAAGTATATGTTGTTGTTGATGACACACTCTAGTGTGATCTTAATAATTAATTAATTAAT 1605
QY 1681 AATAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1740
Db CATTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1665
QY 1741 ATTGAAGACCTGCTCATACAGAGAGAACTTGGTTATTTAACAAGTCAAGGCGCTTTA 1800
Db GTTCAAGGACCTGCTCATACAGAGAGAGAACTTGAATTAATTAATTAATTAATTAATTAAT 1719
QY 1801 GAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1860
Db TGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1779
QY 1861 ACAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1920
Db TCAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1836
QY 1921 ATACCACTCAAGACTCAACAACAATTTTCTGTCATTAATTAATTAATTAATTAATTAAT 1980
Db GAACTGGTATGACATCAACCACTTTTCTGTCATTAATTAATTAATTAATTAATTAATTAAT 1896
QY 1897 AAGATTTTCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1956
Db GAAATTTTGGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2040
QY 2041 CCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2097
Db TCTCTGTGTTTATGTTGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2016
QY 2098 GAAATTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2157
Db GAAATTTGCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2017

Db 1263 TTTTATTAATGTCATAGCTAGATTAATTAATCTAAATGATTAATTAATTTAG 1322
Qy 1389 TCAATTAATGATTAATTAATTAATGCTCATTAACAACACTCAATTAATTCAGAGAG 1448
Db 1323 TAAATGATTTTATTAATTAATTAATGATTAATGATTAATTAATTAATTAATTAAT 1382
Qy 1449 GTCT-----TTATCTAATTAATCAAAACAACTTTTTCATTTCTAGAAAAAGA 1502
Db 1383 ATCTGGGCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1442
Qy 1503 CTGCAATCTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1562
Db 1443 AAGAGAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1502
Qy 1563 ATCCCATTTTCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1622
Db 1503 ATC-----ATTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1545
Qy 1623 TACAGGTGTA--TTAGATGACACACAGTGTGTGATTAATTAATTAATTAATTAATTAAT 1680
Db 1546 CAAAGTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1605
Qy 1681 AAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1740
Db 1606 CATTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1665
Qy 1741 ATTGAAGAGCTGTGTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1800
Db 1666 GTTCAAGAGCTGTGTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1719
Qy 1801 GAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1860
Db 1720 AAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1779
Qy 1861 ACAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1920
Db 1780 TCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1836
Qy 1921 ATACCACTCAAGAGCTCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1980
Db 1837 GAACTGGGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1896
Qy 1981 GGAATTTTGGGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2040
Db 1897 AAAGATTTTGGGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1956
Qy 2041 CCATTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2097
Db 1957 TCTCTGT 2016
Qy 2098 GAATTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2157
Db 2017 GAATTTTGGGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2076
Qy 2158 ATCCAAACAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2200
Db 2077 GTACAACAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2119

RESULT 3

US-09-224-024-30
Sequence 30, Application US/09224024
Patent No. 6056953

GENERAL INFORMATION:

APPLICANT: Leslie Hickie
TITLE OF INVENTION: Materials and Methods for the Control of
NUMBER OF SEQUENCES: 31
CORRESPONDENCE ADDRESS:
ADDRESSEE: David R. Salivanchik
STREET: 2421 N.W. 41st Street, Suite A-1
CITY: Gainesville

STATE: FL
COUNTRY: USA
ZIP: 32606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/224,024
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/856,226
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Salivanchik, David R.
REGISTRATION NUMBER: 31,794
REFERENCE/DOCKET NUMBER: MA79
TELEPHONE: 904-375-8100
TELEFAX: 904-372-5800
INFORMATION FOR SEQ. ID NO.: 30:
SEQUENCE CHARACTERISTICS:
LENGTH: 2061 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-09-224-024-30

Query Match 8.5%; Score 188.2; DB 3; Length 2061;
Best Local Similarity 50.0%; Pred. No. 1.6e-28;
Matches 1107; Conservative 0; Mismatches 923; Indels 185; Gaps 18;

Qy 1 ATGAATCAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 60
Db 1 ATGAATCAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 54
Qy 61 TTTCGGAACGAAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 120
Db 55 GGTTTAGCAAGCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 114
Qy 121 TTACAAAACAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 180
Db 115 CTGAAAAACGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 174
Qy 181 GATTAATTCAGACATTTGCTAGTGTGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 240
Db 175 AATTAATTCGGGGAATTTGTTAGTCTGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 234
Qy 241 GTATCCGTACTCTGTAGCCGGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 300
Db 235 GTATGAGAACTATGTTAGAGCTTTGCTGCCCT----- 270
Qy 301 ATTAATGCTCTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 360
Db 271 GTCTTACTGAGGTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 330
Qy 361 GAACAAACAAACAGTATGACACAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 420
Db 331 GACCTG---CAATGTTTGGCAGAGATTGTTAAACATCGAGAGAAAGCCCTATTAACAAGAA 387
Qy 421 CCGTTAACAGAAAGCATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 480
Db 388 ATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 441
Qy 481 TTCAAGCTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 540
Db 442 CAATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 501
Qy 541 GAATTAACACATCATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 600

Db 502 GCTAAGCAGTA - CAGTATCTCTTACTACTAGAACCTATATATAGATTAAGATTTAGA 560
Qy 601 AATGTTCAATGATTTTATTCGAGAAATACCGGTTTCCACTGAACTTATATAAGC 660
Db 561 TATGTTAAAA-----ATATGCTAGCTATGAAAT 591
Qy 661 CTATTAATCACTTATTTATGCGAGCTGCTAAATTTTCAATTTATTTATTAACAAGGT 720
Db 592 CCAACACTCCCTGCAATGACAAATAGCTACTGGCACTTGATTTATTAACAATGCT 651
Qy 721 GGTGAATTTGGCTGATGATGGAATGCAATATACCTCTTCAAAATGCACTTATGCT 780
Db 652 GCTACCTATTAACAATATATGCTGCAAAATCAAGGAT-----AAATCCAGTCTTTC 705
Qy 781 GGAACATGATGATCTATATATACTTTAAAAAGAAATATACCTAATATATGATCTAT 840
Db 706 AATTAATCTATTAATCTATGAGGCTATTTAAACGTAAATATCAAGATATATCTGACTAT 765
Qy 841 TGTGCAATACCTATATGAAACAGAACTAAAAATCTTAGAGCAAAACCAATATGAAATG 900
Db 766 TGTATACAAAGTACAAATGACAGACTAATGATGATTAAGAACTAATATAGCAACATGG 825
Qy 901 AGTATATTTAATGACTATGAAAGATATATGACCAATTAATGATATCAATCTCTCA 960
Db 826 AATATGATATATATCTTACCGTTAGAAATGACTTAACTGTGTGATGATCTTATGCTAT 885
Qy 961 TTTCTTATATGATATATAAAGATATAGATTTCAATAGAGAAATAGAAATTAAGGC 1020
Db 886 TTTCAAAATTAATGACCCGAAAAAATATCC-----AATAGGA 921
Qy 1021 AATTAAGAAATGAACTGCAAGAAATTTTATACACTGA -AATAAATTTGATGCTTCC 1079
Db 922 GTTAATCTGAATCTACAGAGAAATTTATGAAATTTATTCAGATCAATTTAGAAC 981
Qy 1080 TCAACTTATGATTCACCAATCTGATGATGATGAAATTTATTAACAGTCAAGTT 1139
Db 982 AATACAGAACTAGAAATGATTAATAGAAATCTTAACTTATTTATGATGATTAACCA 1041
Qy 1140 TAAATTAATTTTCAATTTTGAACAATTTATTTTATACAGAAATTAACAATTTGGGAA 1199
Db 1042 GGGCGTTTATACAGAAATTTCTGAGCAATCTTGAATCTTATGATATTTTCTTT 1101
Qy 1200 TCGTTAGTGGTATTTTCTAATCGTATGACCACTTATTAAGCAATATTAATGAAAC 1259
Db 1102 ACAGATTAACAG -ATGCGCTTTACACATATCTAATGATATGCAATATATCTGGGAGC 1160
Qy 1260 TTTATATGAGAAAGAAACAGTTCAACCAACAAACAAATTAAGCAATTTGAACTT 1319
Db 1161 GGTTCATGCAATTAATTTCTCAAGACATCCAA-----GTAATTTCC 1205
Qy 1320 TAAATTTCAATTTGATGATGATGACATCACTCTGTTTCCCTTATTAACCACTT 1379
Db 1206 TTTTATATGAAACAAACCTATGATTAAGTTCGAAATTTGACAGACATGAGACTACTG 1265
Qy 1380 TATAATTAATCAATTTGAACTTTATTAATGCTCATCTAACAACACTCAATATTT 1439
Db 1266 TATAATTAATGAAATGATTTTTCGAATGAC-----ATGAAATGATTTTCATTTTC 1319
Qy 1440 AGCAGAGAGGCTTATTAATTAATCAAAACAACTTTTTCATTTCTAGAAAAA 1499
Db 1320 ATCCAAATTCACAAATGAAATTAATTAAGAACTGATCTTATATGATTTCAAAACA 1379
Qy 1500 AGACTGCAATCTAGTATTTGATTCAGGTTTCCAAACCTTATTAATGATGATAT 1559
Db 1380 AACATG-----GAAAAATTAAGATTAAGTCAATAC 1409
Qy 1560 TTTATCCATTTTCTATTTACTTATTTCTATGATGATTTGATTAACAGCTCAATAT 1619
Db 1410 TCTATCTATATTAATACTGATATTT-----AATTTCTAGTATTTAGGAAAG 1460
Qy 1620 AGATACAGGTGATTAAGATGACACACAGTATGTTGATTAATGATCAATATCA 1679
Db 1461 AAGAAGAGTTGATTTAGTTGACATATCTATGATTTCCAAATTAATCAATGATTT 1520

Qy 1680 TAAATTAATTAATGATATCCAGCAATCAAGATTAACATCTTGATTAACAATCTTAAAGT 1739
Db 1521 AGATTAATCAATCCCAATCTCAAGCTCTTAATAAGCTTTGAAGTATGATTTCAAAAT 1580
Qy 1740 AATTAAGAGACTGCTGATTAACAGAGAACTTGGTTTATTTTACAAAGTCAAGGCGTTT 1799
Db 1581 TGTGAAGAGTCTGGTACACAGGATGAGACTTGGTAAATTTCTTAAGAATATGATGATTT 1640
Qy 1800 AGAAATTAATGATTAATCTTAATTTCTTACCAATCTTATTTCAATTAAGCTTCAATGCT 1859
Db 1641 TAGAGTATGATTTTAAAAAATGTTTCTGACAAATATCAAGTACGTA---TTGCTATGCT 1697
Qy 1860 TACAATGCTGCTGCAATATCTCTTAATATATCTCTTAACAATCAAGAGATATAG 1919
Db 1698 TACTATGCTCAAAAGACAGCT--ATTTTAACCGGAATATGATATCTATATGATGAG 1755
Qy 1920 AATACCACTCAAGCTCAACACACTTTTCTGATCAAAATTAATTAATTAACAATA 1979
Db 1756 CTCCCTAGTACCACTTCCGCAAAAC-----CAATGCTACAGATTTTAACTA 1805
Qy 1980 CGGATTTTGGTATTTTCCAAATTTCCAGTACAT-----ACATTAACCTTTAA 2030
Db 1806 TGCAGATTTTGGATATGATTAATGATTCAGAAACAGTTCCAAATTAACAATTTGAAGAGA 1865
Qy 2031 TCGAAACATACATTAATTAATTTAATCGTCAAGATGATCAATCAATTTTATCAATTTGA 2090
Db 1866 AGACCTTAATTAATGACTTATATGATGATGATGACCAAAATCAATCATATTAATATATATGA 1925
Qy 2091 TAAATTAATTAATTAATTAATTAATTTCTCTATGACCAAAATGAGAAAAACAAATTT 2150
Db 1926 CAAATTAATTAATTAATTAATTAATTAATTTCTGATGATTAATTAATTAATTAATTAATTA 1985
Qy 2151 AGAACTATCCAAACAAATTAATTAATTTTCACAATTAATTAATTAATTAATTAATTA 2205
Db 1986 AGAAAAACACAGAAATTAATGATTAATTTGTTATTAATTAATTAATTAATTAATTAATTA 2040

RESULT 4
PCT-US94-07902-30
; Sequence 30, Application PC/TUS9407902
; GENERAL INFORMATION:
; APPLICANT:
; APPLICANT: Street address: 4980 Carroll Canyon Road
; APPLICANT: City: San Diego
; APPLICANT: State/Province: California
; APPLICANT: Country: US
; APPLICANT: Postal code/Zip: 92121
; APPLICANT: Phone number: (619) 453-8030 Fax number: (619) 453-6991
; APPLICANT: Telex number:
; TITLE OF INVENTION: Materials and Methods for the Control of
; TITLE OF INVENTION: Calliphoridae Pests
; NUMBER OF SEQUENCES: 31
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: David R. Saliwanchik
; STREET: 2421 N.W. 41st Street, Suite A-1
; CITY: Gainesville
; STATE: FL
; COUNTRY: USA
; ZIP: 32606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US94/07902
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Saliwanchik, David R.
; REGISTRATION NUMBER: 31,794
; REFERENCE/DOCKET NUMBER: MA79

TELECOMMUNICATION INFORMATION:
TELEPHONE: 904-375-8100
TELEFAX: 904-372-5800
INFORMATION FOR SEQ ID NO: 30:
SEQUENCE CHARACTERISTICS:
LENGTH: 2061 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
PCT-US94-07902-30

Query Match 8.5%; Score 188.2; DB 6; Length 2061;
Best Local Similarity 50.0%; Pred. No. 1.6e-28;
Matches 1107; Conservative 0; Mismatches 923; Indels 185; Gaps 18;

1 ATGATATCAAAATTAACGATATATACGATATGAAATTTATGATGCGATACCTCACTTAT 60
1 ATGATATCAATATCAAAATTAAGATGAAATATGAAATATTCATGCTC-----CATCCAT 54
61 TTTCCGAACAGAAACAGTAATGATTTAGATACCTTACACAAATATCCAAATCAACCA 120
55 GATTTTACCAAGCTAAATATACATCTATCTAGATATCCATTAGCAAAATTAAGCCAAATCAACCA 114
121 TTACAAACACAAATTAACAAAGAGCGCTCAATATGCTCAAGGAAATACCAATATGCT 180
115 CTGAAAAACAGAAATTAACAAAGATTTGGCTCAATGCTCAAGATTAATCAACAAATATGCT 174
181 GATAATTTGAGAGACATTTGCTAGTCTGATACAAATTTGCTGCAAGTTAGTGACAGTACTAT 240
175 AATATATGGGGGAAATTTTGTTAGTTCTGAAACATATGTTGAGATTTAGTGACAGTTATAT 234
241 GTATCCGCTACTCTGTTAGCCGGATAGTGCGGCTCACTTCTATATCCGACCAATAGGA 300
235 GTAGTAGCACTATGTTAGAGCTTTGCTGCTCT----- 270
301 AATAAGGCTCATATATATCTTTTGTAACCCATACACGTCTTTTGCGCCGCGGA 360
271 GCTTAGCTGAGGTATATATCTTTTGGAAGCTTTTGCGCATCTTTTGCGCAAGATCT 330
361 GAACAGACAAACAGATATGACACAAATTTATTAATGAGGAATTTTGTGATGACA 420
331 GACCTG---CAATGTTTGGCAGAGATTGTTAAACATCGAGAGAGCCCTATACAGAA 387
421 CCGTTACAGAAAGCATAAACAGCTAAAGTTACAACTTTAGAGAGATTAGCAAAATA 480
388 ATAGATAAAAACATATATATATGTACTACTT-----CTATGTAACACCTATAAAAAT 441
481 TTACAAAGCTATATATACGATATAGATGAGAGAAATTTAAAAAGCTACAGCTCCT 540
442 CAACCTGATTAATATCAAGAAATTTTGCATTAATGAGGACGACGCTACACAGCTAT 501
541 GAATTAACACCATCATCAGCATTTACAAACAGCTGCTTCACTTTAAATATAGATTTAG 600
502 GCTAAGAGAGTA-CATGATCTCTTTACTACCTTAGAACCTATATATGATTAAGATTTTGA 560
601 AATGTTCAAAATGATTTTATTCGAGAAATACCTGTTTCAACTTGAACCTTATPAAAAG 660
561 TATGTTAAAA-----ATAATGCTAGCTATGAAATA 591
661 CTATTTACTACTATTTATGCGCAGAGCTCTAATTTCTATTTAAATTTATACCAAGGT 720
592 CCAACACCTCCCTGATATGACAAATAGCTACTTGGAAGTTTATTTAAAAACATGCT 651
721 GCTGAATTTGGCTGATGATGGAATGACATATACATCTTCACAAAATGAAGCTAATGCT 780
652 GCTACCTATTAACATATATGCTGCAAAATCAAGTAT-----AAATCCAGATCTTTC 705
781 GGAACATCAGATGACTATTTATPAACTTTTAAAAAGAAATATACCTTAATATAGTAAT 840
706 AATTCATCTAATTAATCTATCAGGGCTAATTTAAAAAGTAAATATACAAAGATATAGCATAT 765
841 TGTGCAATTAATCTATATGAAACAGAGCTAAAAATCTTAGAGAGAAACCAATATGAAATGG 900

766 TGTATACAAAGTATCAATGCAAGGCTACTACTATGATATGAACTATATCTAATACCAACATGG 825
901 AGTATATTTAATGATATATGCAAGATATATGACATTAATCTGATATTAATACATCTCTCAA 960
826 AATATGATATATCTTACCGTTTAGAAATGACTCTAATCTGTTAGATCTTATTTGCTATT 885
961 TTTTCTTATATGATATATAAAAAGATATAGAGATTCATAGAGGAATAGAAATGAAAGGCT 1020
886 TTTCCAAATTAATGACCCGAAAAATATCC-----AATGGA 921
1021 ATTAGAAATGATCCACAGAGAAATTTATACAACTGA-AATAATTTTGATGCTCTCC 1079
922 GTTAAATCTGAATCTACAGAGAAATTTATACGAATTTTATTCAGATACATTTAGAAC 981
1080 TCACTTAGAGTTCAACCCATCTAGCTAGAGATGAAATATATTTTAAACAGTGCAGATT 1139
982 ATACAGAACTAGAAAAATGATTTAATACGAAATCTATATTTTACTTGATTAACCA 1041
1140 TAAATTAATTTCAATTTTATAGAACATTTATTTTATACAGAAATATCAAAATTTGGGAA 1199
1042 GGGCGTTTATACAAAGAAATTTCCGAGACATTTGATCTTATGATATATTTTCTTTT 1101
1200 TCGTTTATGTTGATTTCTAATGCTGATGCACTTATAGCAATCTAATCTAATCTGAAC 1259
1102 ACAGTAAACAG-ATGGCTTTTACACATCTAATATGATATGCAACATTAATCTGGGAGC 1160
1260 TTTATATGAGAAAGAACAGATTTCAACCCACAAACAAACAAATAGAACATTTGATCTTA 1319
1161 GGTTCATGACATATATTTCTCAAGACATCCAA-----GATATTCC 1205
1320 TAAAGTTCAATTTGATATGATATGACATACATCTCTGTTTCCCTTATTAACCAACACT 1379
1206 TTTTATAGAAACAAACCTATGATAGTGAAATTTGACATGAGAGATGACTGACA 1265
1380 TATATATATCAATGATGACCTTATTTTAAATGCTCATCTAACACACACTCAAAATATTC 1439
1266 TATATATATGAAATGATATTTTTCGAATAGC-----AGTAAATATTTGATATTC 1319
1440 AGCAGAGGCTCTTATCTAATATATCAAAACAACTTTTTCATTTCTAGAAAAA 1499
1320 ATCAATTAACAATAGAAATATATTAATAAGACATGATCTTATATGATCCAAACCA 1379
1500 AGACTGCATCTAGTATTTGATCCAGGTTGTTACCAACTTATATATCTATAGTCTAT 1559
1380 AACATG-----GAAAAATAAAGAAATATGTCATAC 1409
1560 TTTATCCCATTTTCAATTTATTTACTTATCTTATGATGATGATTAAGCTAACAAATAT 1619
1410 TCTATCTATATTAATACTGATATAT-----ATATTTCACTAGTTAGAGAAAG 1460
1620 AGATACAGTGTATTAAGATGACACACAGTAGTGTATGATATATATGCAATATGACA 1679
1461 AAGAAAGTGTGATTTAGTTAGTGGACACATACATAGTGTGATTTCCAAATATCAATAGATT 1520
1680 TAAATATATTAATCAATGATCCAGCAATCAAGGTAACATCTTGATCAAACTTAAAGT 1739
1521 AGATTAATATCAACCAATATCCAGCTCTAAAAAGCTTGAAGTATCTATTCGAAAT 1580
1740 AATTAAGAGCTGCTGATATACAGAGAAACTTGTTTATTTCAAAAGTCAAGGCGTT 1799
1581 TGTGAAGGTCCTGCTGATACAGAGTGAAGCTTGATTTCTTAAGATATGATGATTT 1640
1800 AGAAATTAATGATGAACCTCTAATTTTACAACTTATTTGATTTGAGCTGATATGCT 1859
1641 TAGAGTATGATTTTAAAAATGTTCTGCAATATCAAGTACGTA---TTGGTATGCT 1697
1860 TACAAATGCTGCGAAATATCTCTCTAATATATCTCTTAACAATACAGAGATATAGG 1919
1698 TACTTAATGCTCAAGCAACAGT--ATTCTTAACCGGAATAGATATCTAATAGGTGAG 1755
1920 AATTAACCTCAAGACTCAACCAACTTTTCTGATGAGAAATTTATATTAATTTACATA 1979

Db 1756 CTCCTAGTACCATTCCGCCAANAAC-----CAAACTGCTACAGATTAACTA 1805
Qy 1980 CGAGATTTTGGGATTTTCCATTTCAGTACAGT-----ACATTACTTTAA 2030
Db 1806 TGCAATTTTGGATTTGATACATTTCCAGAAACAGTTCCAAATTAACATTTGAAGAGA 1865
Qy 2031 TCGAAATACATTCATTTATTTAATTCGTGAGATGATCAAAATTCATTTAATCAATGA 2090
Db 1866 AGACATTTATTAATGACCTTATTTATGTFACCAATATCATTCATTAATATATATTGA 1925
Qy 2031 TAAATTTGAATTTATACCAATTTACTCTCTATGACCAAAATAGAGAAAAAATAAT 2150
Db 1926 CAAATTCGAATTTATTCATCATCTGATTTAGTTTATACAGAGAACCAAAATAT 1985
Qy 2151 AGAACTATCCAAACAAAATAATACATTTTTCACAAATCATAACAAAACACTT 2205
Db 1986 AGAAAAACACAGAAATATGTAATGATTTATTTGTTAATTAACAAAGTTCTT 2040

RESULT 5
US-10-089-678-2
; Sequence 2, Application US/10089678
; Patent No. 6962977
; GENERAL INFORMATION:
; APPLICANT: ASANO, Shinichiro
; TITLE OF INVENTION: PROTEIN HAVING INSECTICIDAL ACTIVITY, DNA ENCODING THE PROTEIN, A
; FILE REFERENCE: 069821
; CURRENT APPLICATION NUMBER: US/10/089,678
; PRIOR FILING DATE: 2002-04-03
; PRIOR APPLICATION NUMBER: JP 2000-236140
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: PCT/JP01/06660
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 3504
; TYPE: DNA
; ORGANISM: Bacillus thuringiensis
; FEATURE:
; NAME/KEY: exon
; LOCATION: (1)..(3501)
; OTHER INFORMATION:
US-10-089-678-2

Query Match 4.5%; Score 98.4; DB 3; Length 3504;
Best Local Similarity 43.4%; Pred. No. 1.6e-10;
Matches 602; Conservative 0; Mismatches 751; Indels 33; Gaps 2;

Qy 612 TGATTTTATTCAGAAATACCTGCTTTCACACTTGAACTTAATAACGCTATTACTACC 671
Db 531 TGATTTTGTGCTAAATCCATCTTTGCAATATCTGACAGAAATACATTAATATAC 590
Qy 672 TATTTATGCGAAGCTGCTAATTTTCAATTAATTTATTAACAAGGCTGAAATGCG 731
Db 591 AGTGTATGACAGACAGCAATTTTACATTTGCTATTTATACAGATGCTTCCATTTTGG 650
Qy 732 TGATGAATGGAATGAGATATACATCTTTCACAAATGAACTTAATGCTGAACATCAGA 791
Db 651 AGCAGAGTGG-----GGATTCAACACGAGAGAAATTTTC 683
Qy 732 TGACATTTATTAACCTTTTAAAGAAATATACCTAATTTACTAATCTATGTCGAAATAC 851
Db 684 CACATTTTATGATCGTCAGGTGACACGTACCGCCCAATATCTCGAATTTTGTAAAGTG 743
Qy 852 CTATGAGACAGGACTTAAATAATCTTAGAGACAGAACCAATATGAATGAGATATTTTA 911
Db 744 GATATAACCTGCTTATGATTAATTAAGAGTACGATGCAAGTGGCTGAAGATATCA 803
Qy 912 TGACTATGAGAAATATATGCAATTAATCTGATTTAGATTCATCTCTCAATTTTCTTTATA 971
Db 804 CCAATTTCCGAAAGAAATGACATTAATCTGATTTAGATTTAGAGGTATTTTCCAAACTA 863

Qy 972 TGATATTAATAAGATATAGATTCATATAGAGAAATAGAAATTAAGCAATTAAGATGA 1031
Db 864 TGACACAGTATCGATATCAATCGAAACAGGCCCACTTACAGGGAAAGTATACAGA 923
Qy 1032 ACTCAGAGAAATTTTATACACTGAAATTAATTTTGAATGCTCTTCTCAACTTAAGAT 1091
Db 924 TCCATATGATTTTAAACAGAGAAACAGGTGTGATTTTGTGGCGCTTGATCACTTAACAG 983
Qy 1092 TCAACCCAACTAGCTAGATGGAATATTAATTAACAGTCAAGTTTAAATATTTTC 1151
Db 984 TGATATTTCTTTTTCAGAGTGAAGCGCTGTATTTGTTTCAACACCTATTTGATAT 1043
Qy 1152 ATTTTGAACAATTTTATTTTATACAGAAATTAACAATTTGCGAATCGTTTATGTTGG 1211
Db 1044 ACTCAGTAATTAATGATTTTATTAACAAGAGCGGGCTTCCCTGATTAATTAAGGAATA 1103
Qy 1212 TATTTCTAATCGTATGACCTACT-----TATGCAATATCTAATCTGAACCTTTATA 1265
Db 1104 CTTGAATATTTGGGTAGGACATTCATTAATAATTAATAATAGATGCTGATCAGCAAT 1163
Qy 1266 TGGAGAAAGAACAGGTTCCACCAACAAACAAATAGACATTTGAATCTTATAAGT 1325
Db 1164 AGAAGTATTAACGGTACGATTAATCTTCAACAAATCAAGTATTTATGACAAATTA 1223
Qy 1326 TTCAATTTGTAATGATAGACATCACTCTGTTTCCCTATTCACACCACTTTATAT 1385
Db 1224 GGATATCTTTGAGGTTGATGATTAAGGGCGGATTTTACTAATTAATACGACAGATATA 1283
Qy 1386 TAATCAAAATGAATTTTATTTAATGCTCATCTAACAACACTCAAAATATTCAGCAGG 1445
Db 1284 TGAAGTTCGATCGTATGTTTATACATGCTTGAAGAAGATCAGAGATCAGATCGTGG 1343
Qy 1446 AGGCTTTATCTAATTTATTAACAACAACCTTTTTCATTTCTAGAAAAAAGAGCTG 1505
Db 1344 AGGTTTATGTACTCAAAACCAATACATGATGCAATGATGTAACAAATTTACAAATAC 1403
Qy 1506 CAATCTAGTATTTGATTCAGGCTGTTTCCACAACTTTAATTAATCTAATGCTATTTATTC 1565
Db 1404 GATGATGAATTCCTCCAGAGAAATGAGCCACTAGTAGAGGATTAACATAGATTTATC 1463
Qy 1566 CCAATTTTCAATTAATTTACTTATTCCTATGATGATGATTAACGCTACAAATTAATGATAC 1625
Db 1464 TCATATACCTCTTATTTCTTTTCTAAGAAATGCTAGTAGCTGCTGCTAATATGCAATCT 1523
Qy 1626 AGGTATATTAAGATGAGACACAGTATGTTGATGATTAATATGCAATATCAGATTAAT 1685
Db 1524 CCTGTATTTGCTTGGACACATCGAGTGGAGATGTTTCAAAATACAGTTTATTCAGATTA 1583
Qy 1686 AATTCAATGATTCAGCAATCAAGGATACATCTTGATCAAACTCTAAGGTAAATGA 1745
Db 1584 AATTACTAGATTCAGATTTGTAAGGACATATCTTTAGTTTACAGTATCTAATGTTATTA 1643
Qy 1746 AGACCTGTCATACAGAGGAAATCTTGTTATTTTCAAAAGTCAAGGCGTTTGAAT 1805
Db 1644 AGGCTCTGATTTACAGAGCAATATCTTAAAGAACAGTAGTGTCCGTATCTTA 1703
Qy 1806 TACATGTAATCTCTAATTTTACACAACTTTATTTTATTTAGCTTGCATATGCTACAAA 1865
Db 1704 TACTAGTGTCTGTGTAATATCAATTAACAAAGATATGTTGCAAAATACGTTATGTC 1763
Qy 1866 TGGTCTGGAATATCTTTCCAAATATATCTTACATATACAGAGATTAATGAAATAC 1925
Db 1764 TTCTACTACTAATCTTACGATTTTGTATACAAATTTCTGGAATCTGCAATTTACTCTATAAA 1823
Qy 1926 ACTCAACGACTCAACAACTTTTCTGTTACAAATTAATTAATTAATTAACATACGAGA 1985
Db 1824 TGTTAATTAACATGATTAAGGAGATGATTTAACTTAATATGATTTGACTTAGCAAC 1883
Qy 1986 TTTTGG 1991
Db 1884 TATTGG 1889

RESULT 6
US-10-089-678-3
; Sequence 3, Application US/10089678
; Patent No. 6962977
; GENERAL INFORMATION:
; APPLICANT: ASANO, Shinichiro
; TITLE OF INVENTION: PROTEIN HAVING INSECTICIDAL ACTIVITY, DNA ENCODING THE PROTEIN, AND METHOD
; FILE REFERENCE: 068821
; CURRENT APPLICATION NUMBER: US/10/089, 678
; PRIOR FILING DATE: 2002-04-03
; PRIOR APPLICATION NUMBER: JP 2000-236140
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: PCT/JP01/06660
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
; LENGTH: 3690
; TYPE: DNA
; ORGANISM: *Bacillus thuringiensis*
US-10-089-678-3

Query Match 4.5%; Score 98.4; DB 3; Length 3690;

Best Local Similarity 43.4%; Pred. No. 1.6e-10; Matches 602; Conservative 0; Mismatches 751; Indels 33; Gaps 2;

```
OY 612 TGATTTATTCGAGAAATACCTGTTCCAACTTGAACTTATTAAGCTATTACTAC 671
DB 717 TGATTTGTTGTGTAATATCCCATCTTTGCAATATCTGACAGAGAACATTTATTC 776
OY 672 TATTTATCGCAAGCTGCTATTTTCATTTAAATTTATTAACAAGGTGCTGAATGGC 731
DB 777 AGTGATGACACAGCAGCAATTTACATTTGCTATTATTACAGATGCTTCATTTTGG 836
OY 732 TGATGAATGGAATGAGATATACATCTTCAAAATTTGAACCTTATGCTGAACATCGA 791
DB 837 AGCAGAGTGG-----GGATTCACACGAGAGAAATTTTC 869
OY 792 TGACTATTATTAACCTTTTAAAGAAATATATACCTTAATATAGTATCTATTGCGAAATAC 851
DB 870 CACATTTTATGATGCTGAGTGACGTCACGCCCAATCTCGATTTATTTGTTAAAGTG 929
OY 852 CTATGAAACAGACTAAATAATCTTAGAGACGACCAAAATGGAATGAGATATATTTAA 911
DB 930 GTATTAACACTGCTTGAATTAATTAAGTACGAACTGCAAGTGGCTGAAGTATCA 989
OY 912 TGACATGAGAGATATATGACCATTTACTGTTATGATATACATCTCTCAATTTTCTTATA 971
DB 990 CCAATTCGAGAGAGAAATGACATTAAGTATTAATTTAGATTTAGAGCTTATTTCCAAACTA 1049
OY 972 TGATATTAATAAGATATAGATTTCAATAGAGAGATATAGATTAATAAGATTAAGAAATGA 1031
DB 1050 TGACACACAGTACGATTCACATTCGAAACAAACGCCCAACTTACAGCGAGATGATACGA 1109
OY 1032 ACTCAAGAGAAATTTATCAACTGAAATTAATTTGATGCTCTTCTCACTTGAAGT 1091
DB 1110 TCCATATGATATTTAAACAGAGAAACAAAGTGATGATTTTGAGGGCTTGCTGACCTTAACAG 1169
OY 1092 TCAACCCCAATGAGTATGAGATATATTAATTAACAGCTGCAAGTTTAAATTTATTTTC 1151
DB 1170 TGATATTTCTTTTTCAGAGATGCGAAAGCGCTGTAATTTGTTCCACCACTTATTTGATAT 1229
OY 1152 ATTTTGAACAATTTATTTTATACAGAAATACAAATTTGCGGAATCGTTAGTTGG 1211
DB 1230 ACTCAGTGAATATGATTTTATACACAGACCGGGCTTCCCTTGAATATATACGAATA 1289
OY 1212 TATTTCTAATGATGACCTACT-----TATGCAATATCTAATCTGAACCTTTATA 1265
DB 1290 CCTTGAATATTTGGGATGAGACATTTATTAATTAATTAATTAATGAGATGCTCATGAGATT 1349
OY 1266 TGGAGAAAGAACAGTTTACCCACACAAACAAATTAAGACATTTGATCTTATTAAGT 1325
```

```
DB 1350 AGAAGTATTAACGATGATTAATCTTAAACAAATCAAGTATTAATATTAGCAATTA 1409
OY 1326 TTCAATTTGAATGATGATGACATCACTCGTGTTCCTTATTCACCACTTATTAAT 1385
DB 1410 GGAATCTTTCAGGTTGATGATCATTAAGGGGCGATTTTAGTAAATTAACGACAGATTA 1469
OY 1386 TAATCAAAATTAAGCTTATTTAAATGAGCTATTAACACACACTCAATATTTAGCAGG 1445
DB 1470 TGGAGTTCCTGATGCTGATTTTACATGCTTGAACAAATTAACGATCAGGATCTTGG 1529
OY 1446 AGGCTCTTATCTAATTAATCAAAACAACTTTTTCATTTCTTGAATTAAGAAAGCTG 1505
DB 1530 AGGTTTACGACTCAAAACCAATTAACATGCAAGTATGACACAAATTAACAAATAC 1589
OY 1506 CAATCTAATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1565
DB 1590 GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1649
OY 1566 CCATTTTCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1625
DB 1650 TCATATCACTCTTATTTCTTTTCTTAAGAAATGCTAGTATGCTGCTAGATATGCAATCT 1709
OY 1626 AGGTATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1685
DB 1710 CCTGATTTTCTTGGACACATGCGAGTGGAGATTTTCAAAATTAACATTTATTCAGATTA 1769
OY 1686 AATTAACATATATCCCAACATTAACAAAGTAACTTGTATACAACTCTAAGTATTTGA 1745
DB 1770 AATTAACATATATCCCAACATTAACAAAGTAACTTGTATACAACTCTAAGTATTTGA 1829
OY 1746 AGAAGCTGCTATTAACAGAGAGAACTTGTATTAATTAACAAAGTAAAGGCGTTTGAAT 1805
DB 1830 AGGTCTGATTAATTAACAGAGAGAACTTGTATTAATTAACAAAGTAAAGGCGTTTGAAT 1889
OY 1806 TACATGTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1865
DB 1890 TACTAGTCTCTGTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1949
OY 1866 TGGTCTGTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1925
DB 1950 TTTCTACTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2009
OY 1926 ACTCAAGAGCTCAACCAACTTTTCTGATTAATTAATTAATTAATTAATTAATTAATTA 1985
DB 2010 TGTTAATTAACATGATTAATTAAGGAGATGATTAATTAATTAATTAATTAATTAATTA 2069
OY 1986 TTTTGG 1991
DB 2070 TATTTGG 2075
```

RESULT 7
US-08-349-867-22
; Sequence 22, Application US/08349867
; Patent No. 5508264
; GENERAL INFORMATION:
; APPLICANT: Bradfisch, Gregory A.
; APPLICANT: Thompson, Mark
; APPLICANT: Schwab, George R.
; TITLE OF INVENTION: No. 5508264el Pesticidal Compositions
; NUMBER OF SEQUENCES: 34
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: David R. Saliwanchik
; STREET: 2421 N.W. 41st Street, Suite A-1
; CITY: Gainesville
; STATE: FL
; COUNTRY: USA
; ZIP: 32606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/349,867
FILING DATE:
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Saliwanchik, David R.
REGISTRATION NUMBER: 31,794
REFERENCE/DOCKET NUMBER: MA86
TELECOMMUNICATION INFORMATION:
TELEPHONE: 904-375-8100
TELEFAX: 904-372-5800
INFORMATION FOR SEQ ID NO: 22:
SEQUENCE CHARACTERISTICS:
LENGTH: 3444 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-349-867-22

Query Match 3.9%; Score 85.8; DB 2; Length 3444;

Best Local Similarity 53.8%; Pred. No. 5.1e-08;
Matches 220; Conservative 0; Mismatches 162; Indels 27; Gaps 1;

QY 589 ATACGATTGGAGAAATGTCACAAATGATTTATTCGAGAAATACCTGGTTCCAACTTGAA 648
DB 388 ATTCGATTGGCTAATACAGACAGCGCTTAATACGCAATTAATTTACACTTACA 447
QY 649 ACTTATAAAACGCTATTACTTACTATTATATGCGCAAGCTGCTAATTTCAATTTAAATTTA 708
DB 448 AGTTTGAATCCCTCTTTATCGGTCTATGTTCAAGCGGCGAATTTACATTTACACTA 507
QY 709 TTACAAACAAGGCTGTAATGCTGATGAAATGGAATGCAATATACATCTTCACAAAT 768
DB 508 TTAAGAGCGCTGATCGTTTGGGAGGTTGGGAGCTGGAATATGCTACTGTTAATAT 567
QY 769 GAACCTAATGCTGAACATCAGATGACTTATTAACCTTTTAAAGAAATATACCTTAA 828
DB 568 C-----ATTATATAGATTAAATCTTATTATGATA 600
QY 829 TATAGTAATCTATTTGCAAAATACCTATAGAACAGACTTAAATAATCTTAGAGACGAACA 888
DB 601 TATACGAACATTTGTTGGACATACATCAAGATTAGAAACTTAGAAGCTACTAAT 660
QY 889 AATATGAATGAGATATTTTATGACTATCGAAGATATATACCACTTACTATTAAGAT 948
DB 661 ACTGACAAATGGCAAGATTCAATCAGTTAGAGAGATTTAACACTTACTGTATTAGAT 720
QY 949 ACCATCTCTCAATTTTCTTTATATAGATATTAATAAGATATAGAGATTCAA 997
DB 721 ATCGTGTCTTTTCCGAACTACAGATGTTAGAACTATCCAAATTCAAA 769

RESULT 8
US-08-349-867-26

; Sequence 26, Application US/08349867
; Patent No. 5508264
; GENERAL INFORMATION:
; APPLICANT: Bradfisch, Gregory A.
; APPLICANT: Thompson, Mark
; APPLICANT: Schwab, George E.
; TITLE OF INVENTION: No. 5508264el Pesticidal Compositions
; NUMBER OF SEQUENCES: 34
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: David R. Saliwanchik
; STREET: 2421 N.W. 41st Street, Suite A-1
; CITY: Gainesville
; STATE: FL
; COUNTRY: USA
; ZIP: 32606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/349,867
FILING DATE:
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Saliwanchik, David R.
REGISTRATION NUMBER: 31,794
REFERENCE/DOCKET NUMBER: MA86
TELECOMMUNICATION INFORMATION:
TELEPHONE: 904-375-8100
TELEFAX: 904-372-5800
INFORMATION FOR SEQ ID NO: 26:
SEQUENCE CHARACTERISTICS:
LENGTH: 3444 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-349-867-26

Query Match 3.9%; Score 85.8; DB 2; Length 3444;

Best Local Similarity 53.8%; Pred. No. 5.1e-08;
Matches 220; Conservative 0; Mismatches 162; Indels 27; Gaps 1;

QY 589 ATACGATTGGAGAAATGTCACAAATGATTTATTCGAGAAATACCTGGTTCCAACTTGAA 648
DB 388 ATTCGATTGGCTAATACAGACAGCGCTTAATACGCAATTAATTTACACTTACA 447
QY 649 ACTTATAAAACGCTATTACTTACTATTATATGCGCAAGCTGCTAATTTCAATTTAAATTTA 708
DB 448 AGTTTGAATCCCTCTTTATCGGTCTATGTTCAAGCGGCGAATTTACATTTACACTA 507
QY 709 TTACAAACAAGGCTGTAATGCTGATGAAATGGAATGCAATATACATCTTCACAAAT 768
DB 508 TTAAGAGCGCTGATCGTTTGGGAGGTTGGGAGCTGGAATATGCTACTGTTAATAT 567
QY 769 GAACCTAATGCTGAACATCAGATGACTTATTAACCTTTTAAAGAAATATACCTTAA 828
DB 568 C-----ATTATATAGATTAAATCTTATTATGATA 600
QY 829 TATAGTAATCTATTTGCAAAATACCTATAGAACAGACTTAAATAATCTTAGAGACGAACA 888
DB 601 TATACGAACATTTGTTGGACATACATCAAGATTAGAAACTTAGAAGCTACTAAT 660
QY 889 AATATGAATGAGATATTTTATGACTATCGAAGATATATACCACTTACTATTAAGAT 948
DB 661 ACTGACAAATGGCAAGATTCAATCAGTTAGAGAGATTTAACACTTACTGTATTAGAT 720
QY 949 ACCATCTCTCAATTTTCTTTATATAGATATTAATAAGATATAGAGATTCAA 997
DB 721 ATCGTGTCTTTTCCGAACTACAGATGTTAGAACTATCCAAATTCAAA 769

RESULT 9
US-08-239-476-22

; Sequence 22, Application US/08239476
; Patent No. 5527883
; GENERAL INFORMATION:
; APPLICANT: Thompson, Mark
; APPLICANT: Schwab, George E.
; TITLE OF INVENTION: Improvement of Delta-Endotoxin Expression in
; NUMBER OF SEQUENCES: 34
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: David R. Saliwanchik
; STREET: 2421 N.W. 41st Street, Suite A-1
; CITY: Gainesville
; STATE: Florida
; COUNTRY: USA
; ZIP: 32606

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/239,476
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Saliwanchik, David R.
REGISTRATION NUMBER: 31,794
REFERENCE/DOCKET NUMBER: MA83
TELECOMMUNICATION INFORMATION:
TELEPHONE: (904) 375-8100
TELEFAX: (904) 372-5800
INFORMATION FOR SEQ ID NO: 22:
SEQUENCE CHARACTERISTICS:
LENGTH: 3444 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-239-476-22

Query Match 3.9%; Score 85.8; DB 2; Length 3444;
Best Local Similarity 53.8%; Pred. No. 5.1e-08;
Matches 220; Conservative 0; Mismatches 162; Indels 27; Gaps 1;

```
QY 589 ATACGATTGGAAGATGTTACAGATGATTTTATGAGAAATACCGTTCCTCACTGAA 648
DB 388 ATTCGATTGCTAATACAGACGACGCTTTAATACAGCAATTAATTTTACACTTACA 447
QY 649 ACTATATAAGCTATTAATCTATTTATGCGCAAGCTGCTAAATTTTCAATTAATTTA 708
DB 448 AGTTTGAATCCCTCTTTATCGGTCTATGTTCAAGCGGGAATTTACATTTATCTA 507
QY 709 TTACAACAAGTCTGAATTTGGCTGATGAATGCAATATACATCTTCCACAAATT 768
DB 508 TTAAGAGACGCTGATCCCTTTGCGCAGGTTGGGACTGATATAGCTACTGTTAATAT 567
QY 769 GAACCTAATGCTGGAACATCAGATGACTATTAATTAACCTTTAAAGAAATATACCTAAA 828
DB 568 C-----ATTATATATAGATTAAATTAATTAATTAATTAATTAATTAATTA 600
QY 829 TATAGTAATATGTCGAATACCTATAGAACAGGACTAAATAATCTTAGAGAGAACCA 888
DB 601 TATAGAAACATTTGTTGGACACATACATCAAGATTTAGAAAATTAAGAGTACTAAT 660
QY 889 AATATGAATGAGATATATTTAATGACTATGAGAGATATATGACCATTAATCTGATTAAT 948
DB 661 ACTGACAAATGGGCAAGATCAATCAGTTTAGAGAGATTTTAAACCTTACTGATTAAT 720
QY 949 ACCATCTCTCAATTTCTTTATATGATATTAATAAGATATAGAGATTCAA 997
DB 721 ATCGTGCTCTTTTTCGAACTACAGATGTAGAACATATCCAAATTTCAA 769
```

RESULT 10
US-08-239-476-26
Sequence 26, Application US/08239476
Patent No. 5527883
GENERAL INFORMATION:
APPLICANT: Thompson, Mark
APPLICANT: Schwab, George E.
TITLE OF INVENTION: Improvement of Delta-Endotoxin Expression in
NUMBER OF SEQUENCES: 34
CORRESPONDENCE ADDRESS:
ADDRESSEE: David R. Saliwanchik
STREET: 2421 N.W. 41st Street, Suite A-1
CITY: Gainesville
STATE: Florida

COUNTRY: USA
ZIP: 32606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/239,476
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Saliwanchik, David R.
REGISTRATION NUMBER: 31,794
REFERENCE/DOCKET NUMBER: MA83
TELECOMMUNICATION INFORMATION:
TELEPHONE: (904) 375-8100
TELEFAX: (904) 372-5800
INFORMATION FOR SEQ ID NO: 26:
SEQUENCE CHARACTERISTICS:
LENGTH: 3444 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-239-476-26

Query Match 3.9%; Score 85.8; DB 2; Length 3444;
Best Local Similarity 53.8%; Pred. No. 5.1e-08;
Matches 220; Conservative 0; Mismatches 162; Indels 27; Gaps 1;

```
QY 589 ATACGATTGGAAGATGTTACAGATGATTTTATGAGAAATACCGTTCCTCACTGAA 648
DB 388 ATTCGATTGCTAATACAGACGACGCTTTAATACAGCAATTAATTTTACACTTACA 447
QY 649 ACTATATAAGCTATTAATCTATTTATGCGCAAGCTGCTAAATTTTCAATTAATTTA 708
DB 448 AGTTTGAATCCCTCTTTATCGGTCTATGTTCAAGCGGGAATTTACATTTATCTA 507
QY 709 TTACAACAAGTCTGAATTTGGCTGATGAATGCAATATACATCTTCCACAAATT 768
DB 508 TTAAGAGACGCTGATCCCTTTGCGCAGGTTGGGACTGATATAGCTACTGTTAATAT 567
QY 769 GAACCTAATGCTGGAACATCAGATGACTATTAATTAACCTTTAAAGAAATATACCTAAA 828
DB 568 C-----ATTATATATAGATTAAATTAATTAATTAATTAATTAATTAATTA 600
QY 829 TATAGTAATATGTCGAATACCTATAGAACAGGACTAAATAATCTTAGAGAGAACCA 888
DB 601 TATAGAAACATTTGTTGGACACATACATCAAGATTTAGAAAATTAAGAGGTAAT 660
QY 889 AATATGAATGAGATATATTTAATGACTATGAGAGATATATGACCATTAATCTGATTAAT 948
DB 661 ACTGACAAATGGGCAAGATCAATCAGTTTAGAGAGATTTTAAACCTTACTGATTAAT 720
QY 949 ACCATCTCTCAATTTCTTTATATGATATTAATAAGATATAGAGATTCAA 997
DB 721 ATCGTGCTCTTTTTCGAACTACAGATGTAGAACATATCCAAATTTCAA 769
```

RESULT 11
US-08-598-305A-22
Sequence 22, Application US/08598305A
Patent No. 5827514
GENERAL INFORMATION:
APPLICANT: BRADFISCH, Gregory A.
APPLICANT: THOMPSON, Mark
APPLICANT: SCHWAB, George E.
TITLE OF INVENTION: No. 5827514el Pesticidal Compositions
NUMBER OF SEQUENCES: 38
CORRESPONDENCE ADDRESS:
ADDRESSEE: Saliwanchik, Lloyd & Saliwanchik
STREET: 2421 N.W. 41st Street, Suite A-1

CITY: Gainesville
STATE: FL
COUNTRY: USA
ZIP: 32606-6669
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/598,305A
FILING DATE: 08-FEB-1996
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/349,867
FILING DATE: 06-DEC-1994
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Sanders, Jay M.
REGISTRATION NUMBER: 39,355
REFERENCE/DOCKET NUMBER: MA86.D1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 352-375-8100
TELEFAX: 352-372-5800
INFORMATION FOR SEQ ID NO: 22:
SEQUENCE CHARACTERISTICS:
LENGTH: 3444 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-598-305A-22

Query Match 3.9%; Score 85.8; DB 2; Length 3444;
Best Local Similarity 53.8%; Pred. No. 5.1e-08;
Matches 220; Conservative 0; Mismatches 162; Indels 27; Gaps 1;
QY 589 ATACGATTGGAATGTTACAAATGATTTTATTCGAGAAATACCTGTTTCCAACTGAA 648
DB 388 ATTGATTTTGTCTAATACAGACAGCGCTTTAATACAGCAATTAATTTTACACTTACA 447
QY 649 ACTTATAAAGCGCTATTACTACTATTATTAATGCGAAGCTGCTAATTTTCATTAAATTTA 708
DB 448 AGTTTGAATCCCTCTTTATCGGTCTATGTTCAAGCGGCAATTTACATTATCACCTA 507
QY 709 TTACAACAAGTGTGCTGATTTGCTGATGAATGCAATATACATCCCTTCACAAA 768
DB 508 TTAAGAGAGCGCTGATCGTTTGGGAGGTTGGGAGCTGATATGCTACTGTTATATAT 567
QY 769 GAACCTAATGCTGGAACATCAGATGACTATTATTAATCTTTAAAGAAATATACCTAA 828
DB 568 C-----ATTATATAGATTATATTAATCTTATATATGA 600
QY 829 TATAGTAATCTATTGCAAAATACCTATAGAACAGACTTAAATCTTAGAGACGACCA 888
DB 601 TATAGAAACATTTGTTGGACACATACATCAAGATTAGAAAATTAAAGAGTACTAAT 660
QY 889 AATATGAATGAGATATTTTATATGACTATGCAAGATATATGACCATTAAGTATTAAGT 948
DB 661 ACTGACAAATGGGCAAGATTCATCATGATTAGAGAGATTTAAACCTTACTGTTATTAAGT 720
QY 949 ACCATCTCAATTTTCTTTATATGATATTAATAAAGATATAGATTCAA 997
DB 721 ATCGTGTCTTTTTCGAACTACATGTTAGAACATATCCATTCAA 769

RESULT 12
US-08-598-305A-26
Sequence 26, Application US/08598305A
Patent No. 5827514
GENERAL INFORMATION:
APPLICANT: BRADFISCH, Gregory A.
APPLICANT: THOMPSON, Mark

APPLICANT: SCHWAB, George E.
TITLE OF INVENTION: No. 5827514e1 Pesticidal Compositions
NUMBER OF SEQUENCES: 38
CORRESPONDENCE ADDRESS:
ADDRESSEE: Saliwanchik, Lloyd & Saliwanchik
STREET: 2421 N.W. 41st Street, Suite A-1
CITY: Gainesville
STATE: FL
COUNTRY: USA
ZIP: 32606-6669
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/598,305A
FILING DATE: 08-FEB-1996
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/349,867
FILING DATE: 06-DEC-1994
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Sanders, Jay M.
REGISTRATION NUMBER: 39,355
REFERENCE/DOCKET NUMBER: MA86.D1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 352-375-8100
TELEFAX: 352-372-5800
INFORMATION FOR SEQ ID NO: 26:
SEQUENCE CHARACTERISTICS:
LENGTH: 3444 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-598-305A-26

Query Match 3.9%; Score 85.8; DB 2; Length 3444;
Best Local Similarity 53.8%; Pred. No. 5.1e-08;
Matches 220; Conservative 0; Mismatches 162; Indels 27; Gaps 1;
QY 589 ATACGATTGGAATGTTACAAATGATTTTATTCGAGAAATACCTGTTTCCAACTGAA 648
DB 388 ATTGATTTTGTCTAATACAGACAGCGCTTTAATACAGCAATTAATTTTACACTTACA 447
QY 649 ACTTATAAAGCGCTATTACTACTATTATTAATGCGAAGCTGCTAATTTTCATTAAATTTA 708
DB 448 AGTTTGAATCCCTCTTTATCGGTCTATGTTCAAGCGGCAATTTACATTATCACCTA 507
QY 709 TTACAACAAGTGTGCTGATTTGCTGATGAATGCAATATACATCCCTTCACAAA 768
DB 508 TTAAGAGAGCGCTGATCGTTTGGGAGGTTGGGAGCTGATATGCTACTGTTATATAT 567
QY 769 GAACCTAATGCTGGAACATCAGATGACTATTATTAATCTTTAAAGAAATATACCTAA 828
DB 568 C-----ATTATATAGATTATATTAATCTTATATATGA 600
QY 829 TATAGTAATCTATTGCAAAATACCTATAGAACAGACTTAAATCTTAGAGACGACCA 888
DB 601 TATAGAAACATTTGTTGGACACATACATCAAGATTAGAAAATTAAAGAGTACTAAT 660
QY 889 AATATGAATGAGATATTTTATATGACTATGCAAGATATATGACCATTAAGTATTAAGT 948
DB 661 ACTGACAAATGGGCAAGATTCATCATGATTAGAGAGATTTAAACCTTACTGTTATTAAGT 720
QY 949 ACCATCTCAATTTTCTTTATATGATATTAATAAAGATATAGATTCAA 997
DB 721 ATCGTGTCTTTTTCGAACTACATGTTAGAACATATCCATTCAA 769

RESULT 13

US-08-639-923A-22
Sequence 22, Application US/08639923A
Patent No. 5840554
GENERAL INFORMATION:
APPLICANT: Thompson, Mark
TITLE OF INVENTION: Improvement of Delta-Endotoxin Expression in
NUMBER OF INVENTION: Pseudomonas fluorescens
CORRESPONDENCE ADDRESSES: 38
ADDRESS: David R. Saliwanchik
STREET: 2421 N.W. 41st Street, Suite A-1
CITY: Gainesville
STATE: Florida
COUNTRY: USA
ZIP: 32606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/639,923A
FILING DATE: 24-APR-1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/239,476
FILING DATE: 06-MAY-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Saliwanchik, David R.
REGISTRATION NUMBER: 31,794
REFERENCE/DOCKET NUMBER: MA83.D1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (904) 375-8100
TELEFAX: (904) 372-5800
INFORMATION FOR SEQ ID NO: 22:
SEQUENCE CHARACTERISTICS:
LENGTH: 3444 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-639-923A-22

Query Match 3.9%; Score 85.8; DB 2; Length 3444;
Best Local Similarity 53.8%; Pred. No. 5.1e-08;
Matches 220; Conservative 0; Mismatches 162; Indels 27; Gaps 1;

QY 589 ATACGATTGGAATGTTCCACATGATTTTATTCGAGAAATACCTGGTTCCAACTTGA 648
DB 388 ATTCGATTGCTAATACAGACGCTTTAATACAGCAATTAATTTTACACTTACA 447
QY 649 ACTATATAAAGCGTATTACTACTATTATTCGCGAAGCTGTAATTTTCAATTAATTTA 708
DB 448 AGTTTGAATCCCTCTTTATCGCTATGTTCAGGCGGAATTACATTATTCACCTA 507
QY 709 TTACAACAAGGCTGTAATTCGCTGATGATGATGATGATATACATCTTCCAAATTT 768
DB 508 TTAAGAGACGCTGATGCTTTGGGCAAGGTTGGGACGATATAGCTACTGTTAATAT 567
QY 769 GAACCTAATGCTGGAACATCAGATGACTATTATTAATCTTTTAAAGAAATATACCTAA 828
DB 568 C-----ATTATATAGATTAAATTAATCTTATTCATGCA 600
QY 829 TATAGTACTATTGCGAAATACCTATATGAACAGACTAAATCTTATGAGAGAACCA 888
DB 601 TATACGAAACATGTTGGACACATACATCAAGGATTGAAATCTTAAGGGTACTAAT 660
QY 889 AATATGAATGAGATATTTAATGACTATGAGAGATATATGACATTAATCTGTTATAGAT 948
DB 661 ACTGACATGCGCAAGTTCAATCACTTTAGAGAGATTTTAACACTTACTGTTATAGAT 720

QY 949 ACCATCTGCATTTCTTATATGATATATAAAGATATAGAGTTCAA 997
DB 721 ATCGTGTCTTTTCCGAACGATGTTAGAACATATCAATTCATCAA 769

RESULT 14
US-08-639-923A-26
Sequence 26, Application US/08639923A
Patent No. 5840554
GENERAL INFORMATION:
APPLICANT: Thompson, Mark
TITLE OF INVENTION: Improvement of Delta-Endotoxin Expression in
NUMBER OF INVENTION: Pseudomonas fluorescens
CORRESPONDENCE ADDRESS: 38
ADDRESS: David R. Saliwanchik
STREET: 2421 N.W. 41st Street, Suite A-1
CITY: Gainesville
STATE: Florida
COUNTRY: USA
ZIP: 32606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/639,923A
FILING DATE: 24-APR-1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/239,476
FILING DATE: 06-MAY-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Saliwanchik, David R.
REGISTRATION NUMBER: 31,794
REFERENCE/DOCKET NUMBER: MA83.D1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (904) 375-8100
TELEFAX: (904) 372-5800
INFORMATION FOR SEQ ID NO: 26:
SEQUENCE CHARACTERISTICS:
LENGTH: 3444 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-639-923A-26

Query Match 3.9%; Score 85.8; DB 2; Length 3444;
Best Local Similarity 53.8%; Pred. No. 5.1e-08;
Matches 220; Conservative 0; Mismatches 162; Indels 27; Gaps 1;

QY 589 ATACGATTGGAATGTTCCACATGATTTTATTCGAGAAATACCTGGTTCCAACTTGA 648
DB 388 ATTCGATTGCTAATACAGACGCTTTAATACAGCAATTAATTTTACACTTACA 447
QY 649 ACTATATAAAGCGTATTACTACTATTATTCGCGAAGCTGTAATTTTCAATTAATTTA 708
DB 448 AGTTTGAATCCCTCTTTATCGCTATGTTCAGGCGGAATTACATTATTCACCTA 507
QY 709 TTACAACAAGGCTGTAATTCGCTGATGATGATGATGATATACATCTTCCAAATTT 768
DB 508 TTAAGAGACGCTGATGCTTTGGGCAAGGTTGGGACGATATAGCTACTGTTAATAT 567
QY 769 GAACCTAATGCTGGAACATCAGATGACTATTATTAATCTTTTAAAGAAATATACCTAA 828
DB 568 C-----ATTATATAGATTAAATTAATCTTATTCATGCA 600
QY 829 TATAGTACTATTGCGAAATACCTATATGAACAGACTAAATCTTATGAGAGAACCA 888

Db 601 TATACGAACATGTTGGACACATACATCAAGATTAGAAAAGTTAGAGCTACTAAT 660
Qy 889 AATATGAATGAGATATATTATATGACCTATGACATATATATGACCTATATAGAT 948
Db 661 ACTGACAAATGGGCAAGATTCATCATGTTAGAGAGATTAACTTACCTGATATAGAT 720
Qy 949 ACCATCTCTCAATTTCTTTATATGATATATATAAAGATATAGAGATTCAA 997
Db 721 ATCGTTGCTCTTTTCCGAACCTACGATGTAGAACATATCCAAATTCAAA 769

RESULT 15

PCT-US95-05431-22

Sequence 22, Application PC/TUS9505431

GENERAL INFORMATION:

APPLICANT: Street address: 5501 Oberlin Drive

APPLICANT: City: San Diego

APPLICANT: State/Province: California

APPLICANT: Country: US

APPLICANT: Postal code/Zip: 92121

APPLICANT: Phone number: (619) 453-8030 Fax number: (619) 453-6991

APPLICANT: Telex number:

TITLE OF INVENTION: Improvement of Delta-Endotoxin Expression in

TITLE OF INVENTION: Pseudomonas fluorescens

NUMBER OF SEQUENCES: 34

CORRESPONDENCE ADDRESS:

ADDRESSER: David R. Saliwanchik

STREET: 2421 N.W. 41st Street, Suite A-1

CITY: Gainesville

STATE: Florida

COUNTRY: USA

ZIP: 32606

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: PCT/US95/05431

FILING DATE:

CLASSIFICATION:

ATTORNEY/AGENT INFORMATION:

NAME: Saliwanchik, David R.

REGISTRATION NUMBER: 31,794

REFERENCE/DOCKET NUMBER: M83

TELECOMMUNICATION INFORMATION:

TELEPHONE: (904) 375-8100

TELEFAX: (904) 372-5800

INFORMATION FOR SEQ ID NO: 22:

SEQUENCE CHARACTERISTICS:

LENGTH: 3444 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: DNA (genomic)

PCT-US95-05431-22

Query Match 3.9%; Score 85.8; DB 6; Length 3444;

Best Local Similarity 53.8%; Pred. No. 5.1e-08;

Matches 220; Conservative 0; Mismatches 162; Indels 27; Gaps 1;

Qy 589 ATACGATTTGAGAAATGTTCAATGATTTTATTCGAGAAATACCTGGTTCCAACTTGA 648
Db 388 ATTCGATTTGCTTAATACAGACAGCGCTTAATACAGCAATTAATTTTACACTTACA 447
Qy 649 ACTTATAAACGCTATTACTATTTATGCGCAAGCTGTAATTTTCAATTTAAATTTA 708
Db 448 AGTTTGAATCCCTCTTTATCGCTATGTTCAAGCGCGCAATTTTACATTTATCACAT 507
Qy 709 TTACAACAAGTGCTGAATGGCTGATGATGAATGCAATATACATCTTCACAAATT 768
Db 508 TTAAGAGACGCTGTATCGTTGGGACAGGGTTGGGCACTGATATAGCTACTGTATAAT 567

Qy 769 GAACCTAATGCTGGAACATCAGATGACTATTATTAATCTTTAAAGAAAATATACCTAAA 828
Db 588 C-----ATTATTAATAGATTATTAATCTTATTTATAGA 600
Qy 829 TATAGTAATCTATTTGCAAAATACCTATAGAAACAGGACTPAAAAATCTTAGACAGAACCA 888
Db 601 TATACGAACATTTGTTGGACACATACATCAAGGATTTAGAAAACCTTAAGAGGTACTAAT 660
Qy 889 AATATGAATGAGATATATTATATGACCTATGGAAGATATATGACCATTAATCTTATAGAT 948
Db 661 ACTGACAAATGGGCAAGATTCATCATGTTAGAGAGATTAACTTACCTGATATAGAT 720
Qy 949 ACCATCTCTCAATTTCTTTATATGATATATAAAGATATAGAGATTCAA 997
Db 721 ATCGTTGCTCTTTTCCGAACCTACGATGTAGAACATATCCAAATTCAAA 769

Search completed: December 11, 2005, 10:14:37

Job time : 411 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 11, 2005, 13:37:06 ; Search time 8805 Seconds
(without alignments)
11732.638 Million cell updates/sec

Title: US-10-783-417-1
Perfect score: 2208
Sequence: 1 atgatacaataacagataa.....atcatatacaaacacttga 2208

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 41078325 seqs, 23393541228 residues

Total number of hits satisfying chosen parameters: 82156650

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_est3:*
4: gb_hcc:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_est7:*
9: gb_ges1:*
10: gb_ges2:*
11: gb_ges3:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	94.6	4.3	1392	10	CG757503 P052-4-C0
C 2	90	4.1	1254	10	AG349719 Mus muscu
C 3	88.2	4.0	1758	10	CU509408 SAIL 811
C 4	87.6	4.0	1896	10	CG753083 P048-1-C0
C 5	87.2	3.9	1388	10	AG278124 Mus muscu
C 6	86.8	3.9	1101	10	CNS00396
C 7	84.6	3.8	1348	10	CG749499
C 8	84	3.8	1101	10	CNS006VL
C 9	82.8	3.7	1238	10	AJ925855
C 10	82.4	3.7	1489	10	AG350139 Mus muscu
C 11	82.4	3.7	1594	10	CU110653
C 12	82.2	3.7	1277	9	CC253231
C 13	80.4	3.6	1268	10	AG347098 Mus muscu
C 14	80	3.6	1242	10	CU068807 CH216-115
C 15	79.6	3.6	956	10	CW938578 TC818.3 G
C 16	79.4	3.6	1101	10	AL063921 Drosophila
C 17	78.4	3.6	2157	10	CU081966 CH216-165
C 18	78.2	3.5	1391	10	CG754863 P050-2-G0
C 19	77	3.5	1592	10	CG750135 P044-3-D0
C 20	76.8	3.5	1581	10	CU082000 CH216-165
C 21	76.6	3.5	994	11	CU082000 Tetradon
C 22	76.6	3.5	1325	10	CW952344 TC838.1_D

C 23	76	3.4	1193	10	CG745316 P038-1-G0
C 24	75.6	3.4	1189	1	AJ925925
C 25	75.6	3.4	1227	10	AG430010 Mus muscu
C 26	74.6	3.4	1352	10	AG381852 Mus muscu
C 27	74.4	3.4	734	10	CNS010MP
C 28	74.2	3.4	1101	10	CW942516
C 29	74.2	3.4	1359	8	DN685273
C 30	74	3.4	1092	1	AJ927993
C 31	73.8	3.3	1542	10	AG366981 Mus muscu
C 32	73.6	3.3	1594	10	CU08406
C 33	73.4	3.3	1076	1	AJ926041
C 34	73.2	3.3	1190	10	CNS020N7
C 35	73.2	3.3	1696	10	AG346840 Mus muscu
C 36	73.2	3.3	1928	10	CU073845
C 37	73	3.3	1101	10	CNS00K85
C 38	73	3.3	1378	8	DN736559
C 39	72.8	3.3	1083	1	AJ926215
C 40	72.8	3.3	1260	10	CU491610
C 41	72.6	3.3	1101	10	CNS00EVL
C 42	72.4	3.3	1539	10	AG340947 Mus muscu
C 43	72.2	3.3	1101	10	CNS017V2
C 44	72.2	3.3	1140	10	CU071552
C 45	72.2	3.3	1378	10	AG350209 Mus muscu

ALIGNMENTS

RESULT 1
CG757503/c 1392 bp DNA linear GSS 24-OCT-2003
LOCUS P052-4-C08.za Ppa EcORI BAC library Pristionchus pacificus genomic,
DEFINITION genomic survey sequence.
ACCESSION CG757503
VERSION CG757503.1 GI:37986131
KEYWORDS GSS.
SOURCE Pristionchus pacificus
ORGANISM Pristionchus pacificus
Eukaryota; Metazoa; Nematoda; Chromadorea; Diplogasterida;
Neodiplogasteridae; Pristionchus.
1 (bases 1 to 1392)
Strinivasan,J., Sinz,W., Jeeze,T., Wiggers-Bereballe,L., Jansen,K.,
Buntjer,J., van der Meulen,M. and Sommer,R.J.
An integrated physical and genetic map of the nematode Pristionchus
pacificus
Mol. Genet. Genomics 269 (5), 715-722 (2003)
JOURNAL PUBMED
COMMENT Contact: Sommer RJ
Evolutionary Biology
Max-Planck-Institute for Developmental Biology
Spemannstr. 37-39, Tuebingen D-72076, Germany
Tel: 00497071601371
Fax: 00497071601498
Email: ralf.sommer@tuebingen.mpg.de
Class: BAC ends.

FEATURES

source location/Qualifiers
1..1392
/organism="Pristionchus pacificus"
/mol_type="genomic DNA"
/strain="California"
/db_xref="taxon:54126"
/clone_lib="Ppa EcORI BAC library"
/note="The library was generated by a partial digest of
the genomic DNA with EcORI and cloning into the BAC
vector."

ORIGIN

Query Match 4.3% Score 94.6; DB 10; Length 1392;
Best Local Similarity 44.6% Pred. No. 7.5e+08;
Matches 536; Conservative 0; Mismatches 654; Indels 12; Gaps 4;

360 AGAACAGACAAAACGATGACCAATTTATTAATGGAGAAATTTTGTTCATAC 419

[illegible]

QY	1491	TAGAAAAAAGACTGCATCTAGTATTGATCCAGGTGTCACCAACTTATACTA	1550
Db	230	AAAAAAAAAAAAAAAAATATATATATTTTAAATATAATTAAAAAAATTTAAAAATATA	171
QY	1551	TA	1552
Db	170	TA	169

RESULT 2				
AG349719/c				
LOCUS				
DEFINITION	AG349719	1254 bp	DNA	linear
	Mus musculus molossinus DNA, clone:MSM01-14610.T7, genomic survey			GSS 18-DEC-2004

ACCESSION	AG349719
VERSION	AG349719.1
KEYWORDS	GI:47923029
SOURCE	GSS.
ORGANISM	Mus musculus molossinus (Japanese wild mouse)

[illegible]

TITLE	AUTHORS
Contribution of Asian mouse subspecies <i>Mus musculus molossinus</i> to	Adachi, K., Noguchi, H., Iwagawa, A., Iwazumi, H., Toyokuni, S., and
	Shiroishi, T.
	Adachi, K., Saitou, N., Hattori, M., Sakaki, Y., Moriawaki, K., and
	Toyokuni, S.

genomic constitution of strain C57BL/6J, as defined by BAC-end sequence-SNP analysis
Genome Res. 14 (12), 2439-2447 (2004)

FORNAD	1.53/7902.5
REFERENCE	2 (bases 1 to 1254)
AUTHORS	Hatotori, M., Toyoda, A., Noguchi, H., Kojima, T. and Sakaki, Y.
TITLE	Direct Submission

JOURNAL
Submitted (17-NOV-2003) Masahira Hattori The Institute of Physical
and Chemical Research (RIKEN), Genomic Sciences Center (GSC),
1-7-22 Sutehiro-Chou, Tsurumi-ku, Yokohama, Kanagawa, 230-0045, Japan
(E-mail: hattori@gsc.riken.jp, URL: <http://bgp.gsc.riken.go.jp/>,
Tel: 81-45-503-9111, Fax: 81-45-503-9170)
COMMENT
Clones are derived from the mouse BAC library M5ng01. For BAC

For further information and to ensure the availability of this information, please contact Akiyoshi Abe (abe@rtc.riken.jp), Tsukuba Institute, Bio Resource Center, The Institute of Physical and Chemical Research (RIKEN) 3-1-1 Koyadai, Tsukuba, 305-0074 Japan
phone: 81-298-36-9189, fax: 81-298-36-9199
e-mail: abe@rtc.riken.jp

```

FEATURES
  FRAMES
    Sequencing : T7
  LIBRARY
    Vector      : pBACet3.6
    R.Site 1    : EcoRI
    R.Site 2    : EcoRI
  Location/Qualifiers

```

```

SOURCE
1. 1424
/organism="Mus musculus molossinus"
/mol_type="genomic DNA"
/sub_species="molossinus"
/db_xref="taxon:57486"
/clone="MSMG01-146B10.T7"
/sex="male"
/tissue_type="mixture of kidney and spleen"
/clone_id="MSMG01 Mouse Male BAC Library"

```

[illegible]

[illegible]

RESULT 3	
Locus	CL509408
LOCATION	CU509408 1758 bp DNA linear GSS 01-APR-2004
DEFINITION	SAIL_811.H11.v3 SAIL Collection Arabidopsis thaliana genomic clone SAIL_811.H11.v3, genomic survey sequence.
ACCESSION	CU509408
VERSION	CL509408.1 GI:46006728
KEYWORDS	GSS.
SOURCE	Arabidopsis thaliana (thale cress)
ORGANISM	Arabidopsis thaliana
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophytes; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons; rosids; euroside II; Brassicales; Brassicaceae; Arabidopsie. 1 (bases 1 to 1758) Sessions,A., Burke,E., Presting,G., Aux,G., McElver,J., Patton,D., Dieterich,B., Ho,P., Bacraden,J., Ko,C., Clarke,J.D., Cotton,D., Bullis,D., Snell,J., Mignel,T., Hutchison,D., Kimerly,B., Mitzei,T., Katagiri,F., Glazebrook,J., Law,M. and Goff,S.A. A high-throughput Arabidopsis reverse genetics system Plant Cell 14 (12), 2985-2994 (2002) 12468722
JOURNAL	Contact: Sessions A
PUBMED	Applied Trait Genetics Syngenta Biotechnology Inc. 3054 Cornwallis Rd., Research Triangle Park, NC 27709, USA Email: allen.sessions@syngenta.com ABRC Stock Number GS836276; T-DNA left border flanking sequences of Syngenta Arabidopsis Insertion Library (SAIL) lines are available through the Arabidopsis Biological Resource Center (ABRC). Sequences represent a pool of amplified genomic regions and not single contiguous sequences. Class: TDNA tagged. Location/Qualifiers 1..1758 /organism="Arabidopsis thaliana" /mol_type="genomic DNA" /ecotype="Columbia" /db_xref="taxon:3702" /clone="SAIL_811.H11.v3" /clone_id="SAIL Collection" /note="T-DNA left border sequences were isolated using a modified Tail-PCR strategy"
FEATURES	
source	
ORIGIN	
Query Match	4.0%; Score 88.2; DB 10; Length 1758;
Best Local Similarity	33.0%; Pred. No. 1.3e-06;
Matches	452; Conservative 0; Mismatches 908; Indels 10; Gaps 1,
Oy	382 ACACAATTTTAAATGGAGGAATTTTGTCATCACCGTTAACGAAACATAAAA 441
Db	131 AAAAAAAAAAAAAAANAANAANAANAANAANAANAANAANAANAANAANAAN 190
Oy	442 CAGCTAAAGTTACAACTTAGAGAAGATTAGCAAAATATTACAAAGCTAATTAACGCA 501
Db	191 NNNNNNNNNNNAANAANAANAANAANAANAANAANAANAANAANAANAANAANAANA 250
Oy	502 TTAGATGATTGAGAAATTTAAAGAAGCTACAAGCTCTGGATTACCACCATCATCAGCA 561
Db	251 AAANAANAANNNNNNNNNNNNNNAANAANAANAANAANAANAANAANAANAANAANN 310
Oy	562 TTACACACAGCTGCCTGACTCTTTAAATACGATTGAGAATGTCCACATGATTTATT 621
Db	311 AAAAAAAAAAANAANAANAANAANAANAANAANAANAANAANAANAATATATATATTTTT 370
Oy	622 CGAAGAAATACCTGGTTCCAACTTGAAACCTATATAAAGCATATACACGAATTTAAGCG 681
Db	371 TAATTATTAANAANAANAANAANAANAANAANTNNNNNNNNNNNTNNNTNNNNNNNN 430
Oy	682 CAAGCTCTAATTTTCATTTAAATTTTTCACAACAGCTCTGAATGGCTGATGAATGCG 741
Db	431 AAAAAAAAAAANAANAANAANAANAANAANAANAANAANAANAANAANNNTTAAAAANNA 490
Oy	742 AATGCAGATATACATCCTTCACAAATTTGAACCTATATCTGGAAACATCAGATGACTAATAT 801

[illegible]

RESULT 4					
CG753083/c					
LOCUS	CG753083	1896 bp	DNA	linear	GSS 24-OCT-2003
DEFINITION	P0481-C01.za Ppa ECORI BAC Library <i>Pristionchus pacificus</i> genomic/c.				
	genomic survey sequence.				

[illegible]

```

OY 1338 TGATGACAGTCACCTCTGTTCCCTATTCACACACTTATTAATCAATTGA 1397
DB 1284 AAAAAAAAAATTTATATATATATATATATATATATATATATATATATAT 1225
OY 1338 ACTTATTTAAATGGCTCATCTAACACACTCAATATTCAGCAGAGGGCTTTATC 1457
DB 1224 ATATATTTT-ATATATATATATATATATATATATATATATATATATATAT 1167
OY 1458 TAATATACAAAACAACTTTTTCATTCATTCGTAAGAAAAGCTGAACTGATAT 1517
DB 1166 TAAATATATATATATATATATATATATATATATATATATATATATATATAT 1107
OY 1518 TGATCAGAGTGTGTCACCAACTTTAAATATATATATATATATATATATATAT 1577
DB 1106 AAAAAATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 1047
OY 1578 ATTTACTATTTCTATGTGATTTGATTAACGCTAACATATTTAGATACAGTGTAG 1637
DB 1046 ATATTTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTA 987
OY 1638 ATGACACAGCTAGTGTGATATATATATATATATATATATATATATATATAT 1697
DB 966 ATATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTAT 927
OY 1698 CCCAGCAATCAAGGTAACAATCTTGATATACAACTTAAGTATTTGAAGACCTGTCA 1757
DB 926 AAAAAATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTA 869
OY 1758 TACAGAGGAAACTGTTTATTTTACAAAGTCAGGCGCTTTAGAAATTTACATGTGAAAC 1817
DB 868 AAAAAATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTA 809
OY 1818 TCCATATTTCTACACATCTTATTTATTTATTTATTTATTTATTTATTTATTTA 1877
DB 808 ATATATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTA 753
OY 1878 TACTCTCTATATATATCTCTTACATATCCAGAGTAATATGAACTCAAGCACT 1937
DB 752 AATATATATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTA 694
OY 1938 CAACAACTTTTCTGTGATCAAAATTTATTTATTTATTTATTTATTTATTTA 1995
DB 693 ATATATTTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTA 634
OY 1996 TTCCATTTTCCAGTACGTAACATTACCTTTAAATCCAAATCTCATTTATTTAT 2055
DB 633 TTATTTAAATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTA 574
OY 2056 CGTGAGATGTATCAAAATTTCAATTTTAAATTTGAATTTGAATTTTCAATTTCT 2115
DB 573 -TTAATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTA 515
OY 2116 TCCCTATGACCAAAATATGAGAAAACAAAAATTTGAAGAACTATCCAAACAAATTAAT 2175
DB 514 AAAAAATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTA 455
OY 2176 ACATTTTTCACAAATCATCAAAAACTTT 2206
DB 454 ATATTTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTA 424

```

```

RESULT 5
AG278124/c 1388 bp DNA linear GSS 18-DEC-2004
LOCUS Mus musculus molossinus DNA, clone:MSM01-048F07.77, genomic survey
DEFINITION sequence.
ACCESSION AG278124
VERSION AG278124.1 GI:47851001
KEYWORDS GSS.
SOURCE Mus musculus molossinus (Japanese wild mouse)
ORGANISM Mus musculus molossinus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

```

```

REFERENCE
AUTHORS
1 Abe, K., Noguchi, H., Tagawa, K., Yuzuriha, M., Toyoda, A., Kojima, T., Ezawa, K., Saitou, N., Hattori, M., Sakaki, Y., Moriwa, K. and Shiroishi, T.
TITLE
Contribution of Asian mouse subspecies Mus musculus molossinus to genomic constitution of strain C57BL/6J, as defined by BAC-end sequence-SNP analysis
JOURNAL
PUBMED 15574823
REFERENCE 2 (bases 1 to 1388)
AUTHORS Hattori, M., Toyoda, A., Noguchi, H., Kojima, T. and Sakaki, Y.
TITLE Direct Submission
JOURNAL Submitted (17-NOV-2003) Masahira Hattori, The Institute of Physical and Chemical Research (RIKEN), Genomic Sciences Center (GSC), Japan 1-7-22 Suenhiro-chou, Tsurumi-ku, Yokohama, Kanagawa, 230-0045, Japan (E-mail: hattori@gsc.riken.jp, URL: http://hgp.gsc.riken.go.jp/, Tel: 81-45-503-9111, Fax: 81-45-503-9170)
COMMENT Clones are derived from the mouse BAC library MSM01. For BAC library availability, please contact Kuniya Abe (abe@rtc.riken.jp). The Institute of Physical and Chemical Research (RIKEN) 3-1-1 Koyadai, Tsukuba, 305-0074 Japan phone: 81-298-36-9189, fax: 81-298-36-9199 e-mail: abe@rtc.riken.jp
PRIMERS
Sequencing : T7
LIBRARY : pBACe3.6
VECTOR : EcoRI
R.Site 1 : EcoRI
R.Site 2 : EcoRI
FEATURES
source
location/Qualifiers
1..1388
/orfasm="Mus musculus molossinus"
/mol_type="genomic DNA"
/sub_species="molossinus"
/db_xref="taxon:57486"
/clone="MSM01-048F07.77"
/sex="male"
/tissue_type="mixture of kidney and spleen"
/clone_id="MSM01 Mouse Male BAC library"
ORIGIN
Query Match 3.9%; Score 87.2; DB 10; Length 1388;
Best Local Similarity 44.5%; Pred. No. 2e-06;
Matches 441; Conservative 0; Mismatches 539; Indels 11; Gaps 2;
OY 387 ATTATTTAAATGGAGAAATTTTGTGATACACCGTTAACAGAAACATTAACACACT 446
DB 1067 ATATTTATTTAAATATATATTTATTTAAATTTAAATATTTAAAGAAATTTAAATTAAT 1008
OY 447 AAAGTTACAACTTTAGAGAGATTAGCAAAATTTACAAAGCTATTAATACAGACTTGA 506
DB 1007 ATTAATTAATATTAATTTATTTATTTAAATTAATTAATTAATTAATTAATTAATTAAT 948
OY 507 TGATTTGAGAAATTTAAAGAACTACCAAGCTCTCGATTTACACATCATGACGTTTCA 566
DB 947 ATATATATTTATTTATTTAAATGATTTAAAT-AAAAAAATTTATTTATTTATTTAAATTA 889
OY 567 ACAAGTCGCTTGACTTTAAATATGATTTGAGAAATTTGACAAATGATTTTATTTGCGA 626
DB 888 TTAATATTAATTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 829
OY 627 AATACCTGCTTCAACTGAACTTATTAAGCTATTTACTACCTATTTATGCGCAAGC 686
DB 828 AAAAAATTTAAAGATTTAAATGAAAAATTAAGATTAATTTAAATTTAAATTAATTAATTA 769
OY 687 TGCTAATTTTCAATTTATTTATTAACAACAGTGTGGAATTTGCGTGAATGGAATGCG 746
DB 768 TAAACAAATTAATTTAAATTAATTTAAAGAAATTAATTAATTAATTTATTAATTAATTA 709
OY 747 AGATATACATCTTCACAAATTTGAACCTAATGCTGGAACATCAGATGATTAATTAAC 806

```


ACCESSION CG749499.1
 VERSION CG749499.1 GI:37970425
 KEYWORDS GSS
 SOURCE Pristionchus pacificus
 ORGANISM Pristionchus pacificus
 Eukaryota; Metazoa; Nematoda; Chromadorea; Diplogasterida;
 Neodiplogasteridae; Pristionchus.
 1 (bases 1 to 1348)
 Srinivasan,J., Sinz,W., Jesse,T., Wiggers-Perebolte,L., Jansen,K.,
 Buntjer,J., van der Meulen,M. and Sommer,R.J.
 An integrated physical and genetic map of the nematode Pristionchus
 pacificus
 JOURNAL Mol. Genet. Genomics 269 (5), 715-722 (2003)
 PUBMED 12884007
 COMMENT Contact: Sommer RJ
 Evolutionary Biology
 Max-Planck-Institute for Developmental Biology
 Spemannstr. 37-39, Tuebingen D-72076, Germany
 Tel: 00497071601371
 Fax: 00497071601498
 Email: ralf.sommer@uebingen.mpg.de
 Class: BAC ends.

FEATURES
 source
 1..1348
 Location/Qualifiers
 /organism="Pristionchus pacificus"
 /mol_type="genomic DNA"
 /strain="California"
 /db_xref="taxon:54126"
 /clone_lib="Ppa Ecoli BAC library"
 /note="The library was generated by a partial digest of
 the genomic DNA with Ecoli and cloning into the BAC
 vector."

ORIGIN
 Query Match 3.8%; Score 84.6; DB 10; Length 1348;
 Best Local Similarity 43.1%; Pred. No. 6.4e-06;
 Matches 506; Conservative 0; Mismatches 660; Indels 9; Gaps 2;

QY 390 TATTAAATGGAGAAATTTTGTGATACACCGTTAAACAGAAACATTAACAGCTAA 449
 DB 1332 TATTAAATTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1273
 QY 450 GTTCAAACTTGAAGATTAGACAATATTACAACTAATATACAGATTAGATGA 509
 DB 1272 ATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1213
 QY 510 TTGGAGAAATTAATAAGACTACAGCTCTGATTACACCATCATGACATTAACA 569
 DB 1212 TTTAATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1153
 QY 570 AGTCGCTTGACTCTTAATAATTCGATTGAGAACTTCAATGATTTTATTCGAAT 629
 DB 1152 ATTTAAATTAATTTATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1093
 QY 630 ACCGCTTCCAACTGGAACCTTAATAACGCTATTACCTATTATGCGCAAGCTC 689
 DB 1092 ATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1033
 QY 690 TAAATTTCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 749
 DB 1032 AAAATTAATAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 973
 QY 750 TATACATCTTCACAAATTTGAACCTTAATGCTGGAACATCAGATGCTATTAACTTT 809
 DB 972 AAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 913
 QY 810 AAAAGAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 869
 DB 912 AAAAAA-----AAAAAATTAATTAATTAATTAATTAATTAATTAATTAAT 860
 QY 870 AAATCTTAGAGACGAACCAATATGAAATGAGTATATTATGACTATCGAAGATATAT 929

DB 859 AAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 800
 QY 930 GACCATTAAGTATTAAGTATTAAGTATTAAGTATTAAGTATTAAGTATTAAGTAT 989
 DB 799 ATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 740
 QY 990 AGATTCAATAGAGAAATTAAGTATTAAGTATTAAGTATTAAGTATTAAGTATTAAG 1049
 DB 739 AAAATTAATAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 680
 QY 1050 TACAACGAAATTAATTTGATCGTCTTCCCACTTAAGTATTAAGTATTAAGTAT 1109
 DB 679 AAAATTAATAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 620
 QY 1110 GATGAATATTAATTAACAGCTGCAAGTTTAAATTAATTTTCAATTTTGAACAATTTAT 1169
 DB 619 ATATTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 560
 QY 1170 TTTTATTAACAGAAATTAATTAATTTGCGGAATCGTTAGTTGTTTCTTAATCGTATGC 1229
 DB 559 AATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 502
 QY 1230 ACTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1289
 DB 501 AAAATTAATAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 442
 QY 1290 AACAAAAACATTAAGACATTTGAATCTTAATTAAGTTTCAATTTGAATGAATGACATC 1349
 DB 441 TTAATAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 382
 QY 1350 ACTCTGTTTCCCTTAATTAACACCACTTATTAATTAATTAATTAATTAATTAAT 1409
 DB 381 TTTAATAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 322
 QY 1410 TGCTCATTAACACACACTCAATTAATTAATTAATTAATTAATTAATTAATTAAT 1469
 DB 321 ATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 262
 QY 1470 CACAACCTTTTCTTAATTTCTTGAATAAAGATGCAATGATTAATTAATTAATTAAT 1529
 DB 261 AAAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 202
 QY 1530 TTCACAACTTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1564
 DB 201 TAAATAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 167

RESULT 8
 CNS00EVL
 LOCUS
 DEFINITION
 Drosophila melanogaster genome survey sequence T7 end of BAC:
 BACR29B23 of RpC1-98 library from Drosophila melanogaster (fruit
 fly), genomic survey sequence.
 AL069706
 AL069706.1 GI:4949849
 GSS.
 Drosophila melanogaster (fruit fly)
 Drosophila melanogaster
 Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 Ephydroidea; Drosophilidae; Drosophila.
 1 (bases 1 to 1101)
 Genoscope.
 Direct Submission
 Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :
 BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
 - Web : www.genoscope.cns.fr)
 Determination of this BAC-end sequence was carried out as part of a
 collaboration with the Berkeley Drosophila Genome Project (BDGP).
 The BDGP is constructing a physical map of the Drosophila
 melanogaster genome using these BACs. For further information
 please see <http://www.fruitfly.org> The BDGP Drosophila
 melanogaster BAC library was prepared by Kazutoyo Osogawa and

Avron Mammosets in Pietter de Jong's laboratory in the Department of Cancer Genetics at the Roswell Park Cancer Institute in Buffalo, NY. The library is named RPCI-98 and was constructed by partial EcoRI digestion of Drosophila DNA provided by the BDSP from the isogenic strain Y2; cn bw sp, the same strain used for the BDSP's p1 and EST libraries. A more detailed description of the library and how to order individual BAC clones, the entire library, or filters for hybridization from the BACPAC Resource Center can be found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.

```

/location/Qualifiers
1..1101
/organism="Drosophila melanogaster"
/mol_type="genomic DNA"
/db_xref="taxon:7227"
/clone="BACR2983"
/clone_1fb="RPLC-98"
note="end : 17"

```

ORIGIN

Query Match 3.8%; Score 84; DB 10; Length 1101;
Best Local Similarity 34.2%; Pred. No. 8.3e-06;
Matches 191; Conservative 103; Mismatches 263; Indels 1; Gaps 14.

[illegible][illegible]

ORGANISM	REFERENCE
<i>Theileria annulata</i>	1 (bases 1 to 1238)
Eukaryota, Alveolata, Apicomplexa, Piroplasmida, Theileridae	
Thelazia	
Patil, A., Remold, H., Berriman, M., Murphy, L., Yeats, C. A., Weir	

TITLE The genome of the host-cell transforming parasite *Theileria annulata* and a comparison with *T. parva*

JOURNAL Unpublished (2005)

COMMENT Contact: Pain A

```
FEATURES      Location/Qualifiers
source        1. .1238
```

ORIGIN

Query Match	3.8%;	Score 82.8;	DB 1;	length 1238;
Best Local Similarity	44.0%;	Pred. No. 1.4e-05;		
Matches 396;	Conservative 0;	Mismatches 500;	Indels 5;	Gaps 1

Qy	37	AAACAGTATGACACCAATTTTAAATGGAGAAATTTTGTGTACACGTTAACAG	430
Db	114	AATTAATTAATAAATTAAATATAAATAATATATATATTATATATATAAATAATAT	173
Qy	431	AAAGCATTAACACGCTAAAGTTACAACTTTAGAGATTTAGACCAATATTCAAGCT	490
Db	174	AATATTAATTAATAAATAAATAAATAATTAATAAATAAATAAATAAATAAATAA	233
Qy	491	ATATATACAGCATTAAGATGATGGAGAAATTTAAAAAGATCAACAGCTCTGATTAACAC	550
Db	234	ATATTAATAAATTTAATATTATATTTTATTAATAAATAAATAAATAAATAAATAA	293
Qy	551	CATCATCAGCATTAACAACAGCTGCTGACCTTAATAATGCAATTGGAATGTTACA	610
Db	294	TAAATTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA	353
Qy	611	ATGATTTTATTCGAGAAATACCTGCTTCCAATCTGAACTTTAAACGCTATTACTAC	670
Db	354	TTTAAATTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA	413
Qy	671	CTATTATGCGCAAGCTGCTAATTTTCATTTAAATTTATACACAGCTGCTGAATGG	730
Db	414	TBAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA	473
Qy	731	CTGATGAATGGAATGCGATATACATCCTTCACAAAATGAACCTAAATGCTGACATCAG	790
Db	474	ATATTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA	533
Qy	791	ATGACATTTATTAACCTTTTAAAGAAATATACCTAAATTAAGTAATCTATGGCAATA	850
Db	534	ATTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA	593
Qy	851	CCTATTAAGACAGCATTAATAAATCTTAGACAG-----ACCAATATGAATGCAAT	905

[illegible]

RESULT	10
AGS50139/c	
LOCUS	AGS50139
DEFINITION	Mus musculus molossinus DNA, clone:MSMg01-146R14.TU, genomic survey sequence.
ACCESSION	AGS50139
VERSION	AGS50139.1
KEYWORDS	GI:47923449
SOURCE ORGANISM	GSS. Mus musculus molossinus (Japanese wild mouse)
REFERENCE AUTHORS	Mus musculus molossinus Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurgnathi; Muridae; Murinae; Mus. Abe,K., Noguichi,H., Tagawa,K., Yuzuriha,M., Toyoda,A., Kojima,T., Esawa,K., Saitou,N., Hattori,M., Sakaki,Y., Moriwaki,K. and Shiioishi,T. Contribution of Asian mouse subspecies Mus musculus molossinus to genomic constitution of strain C57BL/6J, as defined by BAC-end sequence-SNP analysis Genome Res. 14 (12), 2439-2447 (2004) 15574823 2 (bases 1 to 1489) Hattori,M., Toyoda,A., Noguuchi,H., Kojima,T. and Sakaki,Y. Direct Submission Submitted (17-NOV-2003) Masahira Hattori, The Institute of Physical and Chemical Research (RIKEN), Genomic Sciences Center (GSC); 1-7-22 Suehiro-chou,Tsurumi-Ku, Yokohama, Kanagawa, 230-0045, Japan [E-mail:hattorigsc.riken.jp, URL:http://hgpc.gsc.riken-go.jp/, Tel:+81-45-503-9111, Fax:+81-45-503-9170] Clones are derived from the mouse BAC library MSMg01. For BAC library availability, please contact Kuniya Abe (abe@rtc.riken.jp). Tsukuba Institute, Bio Resource Center, The Institute of Physical and Chemical Research (RIKEN) 3-1-1 Koyadai, Tsukuba, 305-0074 Japan phone: +81-298-36-9189, fax: +81-298-36-9199 e-mail: abe@rtc.riken.jp PRIMERS Sequencing : TJ LIBRARY vector : pBACE3.6

[illegible]

Oy		1642	ACAGCAGCTACTGTTGATAGCATATATATGCAATTATCGATTAATAATTTCAGATGTCCCA	1701
Db		664	-----TAAATTAAATATATATATAATATAATATAATATTATTTATATATATATATATA	613
Oy		1702	GCAATTCAAAGGTAAACAATCTTGATGCACAACTCTAAGTAAATGAAGAAGACTGGTCATACA	1761
Db		612	TTTATATATATATATATATATATATTTAAAAATATAATATATTTAAATATATATATATA	553
Oy		1762	GGAGAAAATTGGTTTANTTTACAAAGTCAAGGCGCTTAGAAATTCATGTGAAACTCCT	1821
Db		552	AATTTATATATATATATAATATATATTTTAAATTTAAATAATATATATATATATAT	493
Oy		1822	AATTCACACATCTTATTTCCATTAGACTGTGATATGCTACAAATGCGCTGGAAATACT	1881
Db		492	AATATATATATATATATATATATATATATTTAAAAAGAAATATATATATATATATATA	433
Oy		1882	CTTCCTAATATATCTCTTACATATACAGAGATATATGCAATCACCTCAAGACTCAAC	1941
Db		432	TTAATATATATATATATATATATATATATTTAAAAATATATATATATATATATATAT	373
Oy		1942	AACACTTTTTCTGGTCAAATTTATATATTTTACATACGAGATTTGGTATTTCCA	2001
Db		372	ATTAATTTTTTTTAAATAAATATATATATATATATATATATATATATATATATAT	313
Oy		2002	TTTCCAGTACAGTATACATTCCTTAAATGGAACATACCACTTATATTTAATGTCGA	2061
Db		312	TAAATATATATATATATATATATATATATTTATATTTTAAAAATATTTTAAATATATA	253
Oy		2062	GATGTATCAAAATTCATATTTTATATCATGTATATAAATTGAAT-TTATACCAATATCTTCTC	2120
Db		252	AAATTAATATATATAAAATTTTATTTTAAATTTAAATAAAATATATTTATATATAATTTATA	193
Oy		2121	TATGCACCAAAATATAGAGAAAAACAATAATTGAAACTATCCAAACAAAATTAATATCA	2178
Db		192	TATATTAATTTTATTTAAAAATATTTTATATATATTAATTTAAAAAATTTATATAATATATCA	135
RESULT 11				
LOCUS	CL110653	1584 bp	DNA	linear GSS 05-JAN-2004
DEFINITION	ISB1-53P23_Sp6.1 ISB1 Xenopus tropicalis genomic clone ISB1-53P23,			
ACCESSION	CL110653			
VERSION	CL110653.1	GI:40604288		
KEYWORDS	GSS.			
SOURCE	Xenopus tropicalis (western clawed frog)			
ORGANISM	Xenopus tropicalis			
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
	Anura; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;			
	Xenopodinae; Xenopus; Silurana.			
REFERENCE	1 (bases 1 to 1594)			
AUTHORS	Kremetzki,C., Carter,J., McPherson,J., Warren,W., Graves,T.,			
	Mardis,E. and Wilson,R.			
TITLE	A physical map of the xenopus tropicalis genome			
JOURNAL	Unpublished (2003)			
COMMENT	Contact: Richard K Wilson Genome Sequencing Center Washington University School of Medicine Email: submis@wustl.edu Insert Length: 7500 Std Error: 0..00 Seq primer: Sp6 ATTTAGGTGACACTATAG Class: BAC ends High quality sequence start: 390 High quality sequence stop: 470. Location/Qualifiers 1..1594 /organism="Xenopus tropicalis" /mol_type="genomic DNA" /db_xref="taxon:8364" /clone="ISB1-53P23" /clone_1id="ISB1" /note="Vector: pBelobAC11; ISB-1 Xenopus tropicalis BAC"			
FEATURES				
Source				

ORIGIN	Library Segment 1"
Query Match	3.7%; Score 82.4; DB 10; Length 1594;
Best Local Similarity	41.8%; Pred. No. 1.7e-05;
Matches 573; Conservative	0; Mismatches 790; Indels 8; Gaps 2;
Qy	362 AACAGACAAACAGTATGACACATTTATTAATAATGGAGAAATTTTGTTGATACAC 421
Db	1414 AAAAAAAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1355
Qy	422 CGTTACGAAAGCATTAACAGCTAAAGTTACAACTTTGAAAGATTTAGACAAATAT 481
Db	1354 TATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1295
Qy	482 TACAAAGCTAATATACAGCATTTAGATGATGGAGAAATTAATAAGACTACAGCTCTG 541
Db	1294 AAAAAAAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1235
Qy	542 GATTACCAACCTCATCAGCATTTACACAGCTGCTTGACCTTTAAATACGATTTGGA 601
Db	1234 TAAAAAATAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1175
Qy	602 ATGTTCAATGATTTTATTGAGAAATACCTGGTTCCAACTTGAACTTATATAAGCG 661
Db	1174 AAAAAAAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1115
Qy	662 TATTACTACTATTATGCGCAGCTGTAATTTTCAATTAATTAATTTATACAGAGGTG 721
Db	1114 AAAAAAAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1055
Qy	722 CTGAATGGCGTGAATGAAATGCAATGATATCATCTTCCAAATTTGAACCTAATGCTG 781
Db	1054 ATAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1001
Qy	782 GAACATGAGATGACTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 841
Db	1000 AAAAAAAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 941
Qy	842 GTGCAATTAACCTATAGAACAGACTAAATAATCTTAGACGAAACCAATATGAAATGA 901
Db	940 TAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 881
Qy	902 GTATATTTAATGACTATGAAATATATGACCATTTACTGATTTGATATCAATCTCTCAAT 961
Db	880 AAAAAAAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 821
Qy	962 TTCTTTATATGATATTAATAAGATATAGATTCATATAGAGAAATGAACTTAAGCA 1022
Db	820 AAAAAAAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 761
Qy	1022 TTAAAGATGAACCTACAGAGAAATTTATACACTGAATTAATTTTGATGCTCTCTC 1081
Db	760 GAAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 701
Qy	1082 AACTTAGAGTTCAACCAATCTAGCTAGATGATATTAATTAACAGCTGCAAGTTTA 1144
Db	700 ATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 641
Qy	1144 AATATATTTCACTTTTGAACAATTTATTTTATACGAAATATACAAATTTCCGGAATC 1201
Db	640 AAAAAAAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 581
Qy	1202 GTTATAGTTGATTTCTAATCGTATGACCTTATATAGCAATATCTTAAGTAACTT 1261
Db	580 AAAAAAAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 521
Qy	1262 TATATGAGAAAGACAGTTTCAACCAACCAATAAGCAATTTGAATCTTATA 1321
Db	520 AAAAAAAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 461
Qy	1322 AAGTTCAATTTGATGATGACATCACTCTGTTTCCCTATTCACCACTTTA 1381

	Dd	1219	ATGAATTGAGAAAATTAAAAAGACTACAGCTCCTGGATTACCAACATCATCATGACATTAC	565
	RESULT 13			
	AG347098/c			
	LOCUS	1268 bp	DNA	linear GSS 18-DEC-2004
	DEFINITION	Mus musculus molossinus DNA, clone:MSMg01-142102.T7, genomic survey sequence.		
	ACCESSION	AG347098.1	GI:47920408	
	VERSION	GSS.		
	KEYWORDS	Mus musculus molossinus (Japanese wild mouse)		
	SOURCE	Mus musculus molossinus		
	ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
	REFERENCE	Abe,K., Noguichi,H., Tagawa,K., Yuzuriha,M., Toyoda,A., Kojima,T., Ezawa,K., Saitou,N., Hattori,M., Sakaki,Y., Moriwaki,K. and Shiroyshi,T.		
	AUTHORS	Contribution of Asian mouse subspecies Mus musculus molossinus to genome constitution of strain C57BL/6J, as defined by BAC-end sequence-SNP analysis		
	TITLE	Genome Res. 14 (12), 2439-2447 (2004)		
	JOURNAL	15574823		
	PUBMED	2 (bases 1 to 1268)		
	REFERENCE	Hattori,M., Toyoda,A., Noguichi,H., Kojima,T. and Sakaki,Y.		
	AUTHORS	Submitted (17-NOV-2003) Maahaira Hattori, The Institute of Physical and Chemical Research (RIKEN), Genomic Sciences Center (GSC); Japan		
	TITLE	Direct Submission		
	JOURNAL	Submitted (17-NOV-2003) Maahaira Hattori, The Institute of Physical and Chemical Research (RIKEN), Genomic Sciences Center (GSC); Japan		
	COMMENT	1-7-22 Shunho-chou,Tsurumi-Ku,Tokohama,Kanagawa,230-0045,Japan (E-mail:hattori@gsc.riken.jp, URL:http://hgdp.gsc.riken.go.jp/, Tel:81-45-503-9111, Fax:81-45-503-9170) Clones are derived from the mouse BAC library MSMg01. For BAC library availability, please contact Kunihya Abe (abe@rtc.riken.jp). Tarkuba institute, Bio Resource Center. The Institute of Physical and Chemical Research (RIKEN) 3-1-1 Koyadai,Tsukuba,305-0074 Japan phone: 81-298-36-9189, fax: 81-298-36-9199 e-mail: abe@rtc.riken.jp PRIMERS Sequencing : T7 LIBRARY Vector : pBACE3.6 R.Site 1 : EcoRI R.Site 2 : EcoRI. Location/Qualifiers 1..1268 /organism="Mus musculus molossinus" /mol_type="genomic DNA" /sub_species="molossinus" /db_xref="taxon:57486" /clone="MSMg01-142102.T7" /sex="male" /tissue.type="mixture of kidney and spleen" /clone_lib="MSMg01 Mouse Male BAC Library"		
	FEATURES			
	source			
	ORIGIN			
	Query Match	3.6%	Score 80.4;	DB 10; Length 1268;
	Best Local Similarity	45.1%;	Pred.No.4.1e-05;	
	Matches 418; Conservative	0; Mismatches 495;	Indels 13; Gaps 3;	
Oy	386	AATTATTAATAATGGAGAAATTTTGTGTATCACCCCTTAACAGAAAGCATTAACAGC	445	
Dd	1061	AATAAAAAAAAAATTATTAATTAATTAATAATTAATTAATTAATTAATTAATTAATTA	1002	
Oy	446	TAAAGTTACAACCTTAGAGAGATTTAACAACAATTTTCAAGAGCTATATACAGATTAG	505	
Dd	1001	TAAATTAATATATAATTTTATTAATAATTAATTAATAAATTAATTAATTAATTTT	942	
Oy	506	ATGATTGAGAAAATTAAAAAGACTACAGCTCCTGGATTACCAACATCATCATGACATTAC	565	

Df		941	AATATTAAAAAAATTAATAAATATTAATAATTAATAAAAAAAAAAATTAATAAAAAAAAAAATTAAT	882
Oy		566	AACAGCGCTGACTTAAATAGCATTTGGAAATGTCCACAATGATTTTATTGCAG	625
Df		881	AAATATATATATATAAATAATTAATAAATTTAA- -AATATAAAAATTTATTTTTTATATATA	824
Oy		626	AAATACCTGGTTCCACTTGAAACTTATATAAGCTATTACTATTTATATGCGCAAG	685
Df		823	AATTTAAATATATATATATAAATAAATAATATAAATAATTAATATATATATATATATA	764
Oy		686	CTGCTAATTTTCATTTAATTTATTACACAGGTGCTGAATGGCTGATGATGGAATG	745
Df		763	AAATTAAT---ATTATATTAATTTAAATTAATTAATTAATTTATATATAAATAATATA	707
Oy		746	CAGATATACATCTTCCAATTAATGACCTAATGCTGGAACATCGATGACTATTATAAC	805
Df		706	AAATTTAAATTAATTAATTTAAATTAATAAATAAATAAATAATTAATTAATTAATATA	647
Oy		806	TTTTTAAAAAATAATATCTTAATATAGTAACTAATGTGCCAATACCATTATGAACAGGAC	865
Df		646	TAAATAATTAATTAATAAATAATTAATAAATAATATATATTAATTAATAAATAATTAAT	587
Oy		866	TAAATAATCTTAGACAGAACCAATATGAATGGAGATTTTATATGACTATGGAAGA-	924
Df		586	ATTAATTAATTAATAAATAAATAAATAAATAAATAATATATATATATATATATATATA	527
Oy		925	-----TATATGACATTACTGTATTAAGATACCATCTCATAATTTCTTATATATGATAT	977
Df		526	AAAAATTAATTAATAAATTTATTAATTAATAATTAATTAATTAATAAATAATTAATTAAT	467
Oy		978	AAAAAGATATAGAGATTCATATGAGAGATATGAAGTAAAGSCATTAAGATGAACTCAC	1033
Df		466	AATTTAATTTTAAAAAATAAATAAATAAATAAAGATTAATAATTAATTTATGATTTAATATA	407
Oy		1038	AAGAGAAATTTATCAACTGAAATAATTTTGATGCTGCTCCACACTTAGAGTTCAACC	1097
Df		406	AAAAATATTTATTTTATTTTNTTAAAAAATAAATAATTAATAATTTTATATAAATA	347
Oy		1098	CAATCTAGCTACGATGGAATTAATTTTAACACGTGCAAGTTTAAATTAATTTTCATTTTT	1157
Df		346	TAAAAAATAAATAATTAATTTTAAAAATTTTATTAATAAATAAATAAATAATTAATTAATTT	287
Oy		1158	AGAACATTTTATTTTATATCAGAAATPACAAATTTGGGAAATGTTAGTTGATATTC	1217
Df		286	TAAATTAATTAATTTTAAATTAATTAATTAATAATTAATTAATTAATTAATTAATAA	227
Oy		1218	TAACTGATGACACCTTACTATAGCAATACTATTAAGTAACTTATATATGAGAAAGAAC	1277
Df		226	AAAAAATTAATTTAATTAATTAATAAATAATTAATTAATTAATTAATTAATTAATTAAT	167
Oy		1278	AGGTTCCACCAACAAAAACAATAA	1303
Df		166	TATTTTATTAATAAATAATTAATTAATAA	141

RESULT 14
 CL068807/c 1242 bp DNA linear GSS 31-DEC-2003

LOCUS CH216-115B3, sp6.1 CH216 xenopus tropicalis genomic clone

DEFINITION CH216-115B3, genomic survey sequence.

ACCESSION CL068807

VERSION CL068807.1 GI:40524720

KEYWORDS GSS.

SOURCE Xenopus tropicalis (western clawed frog)

ORGANISM Xenopus tropicalis
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;
 Xenopodinae; Xenopus; Silurana.
 1 (bases 1 to 1242)
 Kremitzki,C., Carter,J., McPherson,J., Warren,W., Graves,T.,
 Mardis,E. and Wilson,R.
 A physical map of the xenopus tropicalis genome
 Unpublished (2003)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 11, 2005, 10:14:46 ; Search time 1318 Seconds
(without alignments)
1165.130 Million cell updates/sec

Title: US-10-783-417-1

Perfect score: 2208

Sequence: 1 atgatacaataacagataa.....atcatatacaaaacacttga 2208

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 4996997 seqs, 3332346308 residues

Total number of hits satisfying chosen parameters: 9993994

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

N Geneseq_21:*

- 1: geneseqn1980s:*
- 2: geneseqn1980s:*
- 3: geneseqn2000s:*
- 4: geneseqn2001as:*
- 5: geneseqn2001bs:*
- 6: geneseqn2002as:*
- 7: geneseqn2002bs:*
- 8: geneseqn2003as:*
- 9: geneseqn2003bs:*
- 10: geneseqn2003cs:*
- 11: geneseqn2003ds:*
- 12: geneseqn2004as:*
- 13: geneseqn2004bs:*
- 14: geneseqn2005s:*

Pred. NO. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2208	100.0	2208	13	ADR89394 AXMI-006
2	1903.8	86.2	2235	13	ADR89396 AXMI-007
3	1794.4	81.3	2085	13	ADR89398 AXMI-007
4	346.6	15.7	3940	1	AAN93054
5	346.6	15.7	4571	1	AAN93059
6	345	15.6	4934	1	AAN81490
7	340.6	15.4	3543	2	AAN14669
8	340.6	15.4	3543	2	AAO81178
9	188.2	8.5	2061	2	AAO14670
10	188.2	8.5	2061	2	AAO81180
11	178.2	8.1	3756	1	AAN50525
12	163.8	7.4	4896	6	ADQ43974
13	163.8	7.4	4896	10	ADF31301
14	163.8	7.4	4896	10	ADF31306
15	98.4	4.5	3690	6	ABK51132
16	94.6	4.3	15548	8	ABJ34155
17	90.4	4.1	8056	8	ABJ10246
18	89.6	4.1	2130	12	ADP71294
19	89.6	4.1	2246	12	ADP71295

20	86	3.9	4660	14	ADZ70902	AdZ70902 Human mat
21	85.8	3.9	3444	2	AAO5249	AaO5249 CryIF/cry
22	85.8	3.9	3444	2	AAO5251	AaO5251 CryIF/cry
23	85.8	3.9	3444	2	AAO18723	AaO18723 CryIF/cry
24	85.8	3.9	3444	2	AAO18701	AaO18701 CryIF/cry
25	85.8	3.9	3444	2	AAV62080	AAV62080 Plasmid p
26	85.8	3.9	3444	2	AAV62082	AAV62082 Plasmid p
27	85.8	3.9	3450	2	AAO5269	AaO5269 CryIA(c)/
28	85.8	3.9	3450	2	AAO18721	AaO18721 CryIA(c)/
29	85.8	3.9	3450	2	AAV62079	AAV62079 Plasmid p
30	85.8	3.9	3522	2	AAO10182	AaO10182 Lepidopte
31	85.8	3.9	3522	2	AAO47291	AaO47291 Delta end
32	85.8	3.9	3522	2	AAO5250	AaO5250 CryIF/436
33	85.8	3.9	3522	2	AAO5270	AaO5270 CryIF/cry
34	85.8	3.9	3522	2	AAO18722	AaO18722 Codon-rev
35	85.8	3.9	3522	2	AAO18702	AaO18702 CryIF/436
36	85.8	3.9	3522	2	AAV62083	AAV62083 Plasmid p
37	85.8	3.9	3522	2	AAV62081	AAV62081 Plasmid p
38	85.8	3.9	3522	2	AAV60606	AAV60606 Wild-type
39	85.8	3.9	3522	4	AAO50565	AaO50565 B. thurin
40	85.8	3.9	3558	4	AAO04420	AaO04420 B. thurin
41	85.8	3.9	4020	2	AAO14648	AaO14648 cryIF gen
42	82.2	3.7	1524	14	ADZ70897	AdZ70897 Human mat
43	81.2	3.7	6175	6	ABJ33307	ABJ33307 Human 1mm
44	80	3.6	2482	14	ADZ71091	AdZ71091 Human chr
45	78.2	3.5	2131	14	ADZ71009	AdZ71009 Human chr

ALIGNMENTS

RESULT 1	ADZ70902	standard; CDNA, 2208 BP.
ID	ADR89394	
XX	ADR89394;	
AC	18-NOV-2004 (first entry)	
XX		
DT		
XX		
DE	AXMI-006 coding sequence.	
XX		
KW	ss; gene; delta-endotoxin; delta-endotoxin associate polypeptide;	
KW	expression cassette; transformation; transgenic; plant; bacteria;	
KW	lepidoptera; coleoptera; pest; pesticide; resistance;	
KW	pesticidal activity.	
XX		
OS	Bacillus thuringiensis.	
XX		
FH	Key	Location/Qualifiers
FT	CDS	1..2208
FT		/*tag=a
FT		/product="AXMI-006"
PN	WO2004074462-A2.	
XX		
PD	02-SEP-2004.	
XX		
PF	20-FEB-2004; 2004WO-US005829.	
XX		
PR	20-FEB-2003; 2003US-0448632P.	
PR	20-FEB-2003; 2003US-0448633P.	
PR	20-FEB-2003; 2003US-0448797P.	
PR	20-FEB-2003; 2003US-0448806P.	
PR	20-FEB-2003; 2003US-0448810P.	
PR	20-FEB-2003; 2003US-0448812P.	
PR	19-FEB-2004; 2004US-00781979.	
PR	19-FEB-2004; 2004US-00782020.	
PR	19-FEB-2004; 2004US-00782036.	
PR	19-FEB-2004; 2004US-00782141.	
PR	19-FEB-2004; 2004US-00782570.	
PR	19-FEB-2004; 2004US-00783417.	
XX		
PA	(ATHE-) ATHENIX CORP.	

Db	1661	AAATAATTACAAATGATCCGAGCANTCAAAGGTACAACTTGTATCAAACTCTTAAGSTA	1740
Qy	1741	ATTGAAAGGACCTGGTCATACAGAGGAAACTTGGTTATTTTCAAAAGTCAGAGGCGTTTA	1800
Db	1741	ATTGAAAGGACCTGGTCATACAGAGGAAACTTGGTTATTTTCAAAAGTCAGAGGCGTTTA	1800
Qy	1801	GAAATTCATGTGAAACTCTCTTAATTTTACAAATCTTAATTTCACTTGACTTCGATAGCT	1860
Db	1801	GAAATTCATGTGAAACTCTCTTAATTTTCAAAATCTTAATTTCACTTGACTTCGATAGCT	1860
Qy	1861	ACAATGGTGTGGAATAACTCTTCCTTAATATCTCTTCAATACCAAGAGTAATGGA	1920
Db	1861	ACAATGGTGTGGAATAACTCTTCCTTAATATCTCTTCAATACCAAGAGTAATGGA	1920
Qy	1921	ATACCACTCAACGACTCTCAACAAACCTTTTCTGGTACAAATTTAATTAATTCAATAC	1980
Db	1921	ATACCACTCAACGACTCTCAACAAACCTTTTCTGGTACAAATTTAATTAATTCAATAC	1980
Qy	1981	GGAAGTTTGGGATTTTCCAAATTTTCGAAGTACAGTAACATTACCTTTAAATCGAAACATA	2040
Db	1981	GGAAGTTTGGGATTTTCCAAATTTTCGAAGTACAGTAACATTACCTTTAAATCGAAACATA	2040
Qy	2041	CCATTTAATTAAATCGAGAGATGTACAAATTCAAATTTAATCACTGATPAAATTTGAA	2100
Db	2041	CCATTTAATTAAATCGAGAGATGTACAAATTCAAATTTAATCACTGATPAAATTTGAA	2100
Qy	2101	TTTATACCAATTACTTCTCTATGCACCAAATATGAGAAAAACAATAATTGAAACTATC	2160
Db	2101	TTTATACCAATTACTTCTCTATGCACCAAATATGAGAAAAACAATAATTGAAACTATC	2160
Qy	2161	CAAAACAAAATAATCAATTTTTCACAAATCATACAAAAACACTTTTGA 2208	
Db	2161	CAAAACAAAATAATCAATTTTTCACAAATCATACAAAAACACTTTTGA 2208	

19-FEB-2004; 2004US-00782306.
PR 19-FEB-2004; 2004US-00782144.
PR 19-FEB-2004; 2004US-00782570.
PR 19-FEB-2004; 2004US-00783417.
XX
XX
PA (ATHE-) ATHENIX CORP.
XX
XX Carozzi N, Hargies T, Koziel MG, Duck NB, Carr B;
XX
XX WPI; 2004-635574/61.
DR P-PSDB; ADR89397.
XX
XX
PT New isolated delta-endotoxin and delta-endotoxin-associated nucleic acids
PT and polypeptides, useful for killing lepidopteran or coleopteran pests or
PT for producing organisms with pesticide resistance.
XX
XX
PS Claim 1; SEQ ID NO 8; 178bp; English.
XX
XX This sequence encodes an isolated delta-endotoxin. Some of the delta-
CC endotoxin coding sequences of the invention have alternative start
CC codons, producing more than one protein from a single open reading frame.
CC The nucleic acid sequences of the invention are useful in DNA constructs
CC or expression cassettes for transformation and expression in plants and
CC bacteria. The nucleic acids and corresponding polypeptides are useful for
CC killing lepidopteran or coleopteran pests. Compositions containing the
CC delta-endotoxins of the invention, and methods for their production, are
CC useful for the production of organisms with pesticide resistance,
CC specifically bacteria and plants. These organisms are useful for
CC generating altered or improved delta-endotoxin or delta-endotoxin-
CC associated proteins that have pesticidal activity, or for detecting the
CC presence of delta-endotoxin or delta-endotoxin-associated proteins or
CC nucleic acids in products or organisms.
XX
XX Sequence 2235 BP; 861 A; 364 C; 316 G; 694 T; 0 U; 0 Other;

RESULT 2
ADR89396
ID ADR89396 standard; cDNA; 2235 BP.

AC ADR89396;

DT 18-NOV-2004 (first entry)

AXMI-007 coding sequence.

ss; gene; delta-endotoxin; delta-endotoxin associate polypeptide; KW

KM lepidoptera; pest; pesticide; resistance;
expression cassette; transformation; plant; bacteria;

perpetual activity.
XX
XX

XX

FT	1. .2235
key	LOCATIONS/QUALITIES
CDS	

```

/ cag= a
/product= "AXMI-007"

```

XX
XX
XX
/ctandt_exechr- pub:t. :), aa:mec

XX
XX

.ZV-ZOLF/OLOOZOV

XX

[illegible]

PR 20-FEB-2003; 2003US-0448633P.

PR 20-FEB-2003; 2003US-0448806P.

PR 20-FEB-2003; 2003US-0448812P.

PR 19-FEB-2004; 2004US-00782020.

Query Match	86.2%	Score 1903.8	DB 13	Length 2235
Best Local Similarity	92.2%	Pred. No. 0		
Matches 2030; Conservative	0	Mismatches 162	Indels 9	Gaps 2

11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847

[illegible][illegible][illegible]

Year	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022	2023	2024	2025	2026	2027	2028	2029	2030	2031	2032	2033	2034	2035	2036	2037	2038	2039	2040	2041	2042	2043	2044	2045	2046	2047	2048	2049	2050	2051	2052	2053	2054	2055	2056	2057	2058	2059	2060	2061	2062	2063	2064	2065	2066	2067	2068	2069	2070	2071	2072	2073	2074	2075	2076	2077	2078	2079	2080	2081	2082	2083	2084	2085	2086	2087	2088	2089	2090	2091	2092	2093	2094	2095	2096	2097	2098	2099	2100
1990	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022	2023	2024	2025	2026	2027	2028	2029	2030	2031	2032	2033	2034	2035	2036	2037	2038	2039	2040	2041	2042	2043	2044	2045	2046	2047	2048	2049	2050	2051	2052	2053	2054	2055	2056	2057	2058	2059	2060	2061	2062	2063	2064	2065	2066	2067	2068	2069	2070	2071	2072	2073	2074	2075	2076	2077	2078	2079	2080	2081	2082	2083	2084	2085	2086	2087	2088	2089	2090	2091	2092	2093	2094	2095	2096	2097	2098	2099	2100

[illegible]

171 AA

[illegible][illegible][illegible]

Year	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022	2023	2024	2025	2026	2027	2028	2029	2030	2031	2032	2033	2034	2035	2036	2037	2038	2039	2040	2041	2042	2043	2044	2045	2046	2047	2048	2049	2050	2051	2052	2053	2054	2055	2056	2057	2058	2059	2060	2061	2062	2063	2064	2065	2066	2067	2068	2069	2070	2071	2072	2073	2074	2075	2076	2077	2078	2079	2080	2081	2082	2083	2084	2085	2086	2087	2088	2089	2090	2091	2092	2093	2094	2095	2096	2097	2098	2099	2100
1990	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022	2023	2024	2025	2026	2027	2028	2029	2030	2031	2032	2033	2034	2035	2036	2037	2038	2039	2040	2041	2042	2043	2044	2045	2046	2047	2048	2049	2050	2051	2052	2053	2054	2055	2056	2057	2058	2059	2060	2061	2062	2063	2064	2065	2066	2067	2068	2069	2070	2071	2072	2073	2074	2075	2076	2077	2078	2079	2080	2081	2082	2083	2084	2085	2086	2087	2088	2089	2090	2091	2092	2093	2094	2095	2096	2097	2098	2099	2100

.....

100

U
C
E
R
T
A
I
N
L
Y
T
H
E
S
E
P
O
L
Y
M
E
R
S
W
I
L
B
E
C
O
M
P
O
S
I
T
E
D
F
R
O
M
T
H
E
S
E
P
O
L
Y
M
E
R
S
W
I
L
B
E
C
O
M
P
O
S
I
T
E
D
F
R
O
M
T
H
E
S
E
P
O
L
Y
M
E
R
S

100

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
84

100

Db 488 ATAAATACAGCATAGATGATGGAGAAAATTAAAGACTACAGCTCTGGATTACAC 547
Qy 551 CATCATCGCATTTACACACAGCTGCTTGACTCTTTAAATAAGATTTGAGAAATGTCACA 610
Db 548 CATCATCGCATTTACACACAGCTGCTTGACTCTTTAAATAAGATTTGAGAAATGTCACA 607
Qy 611 ATGATTTATTCGGAATATACCTGGTTTCCAACTGGAACCTTAATAAACGCTATTATAC 670
Db 608 ATGATTTATTCGGAATATACCTGGTTTCCAACTGGAACCTTAATAAACGCTATTATAC 667
Qy 671 CTATTTATGCGCAAGCTGCTAATTTTCATTTAAATTTTACCAACAGGCTGTAATGG 730
Db 668 CTATTTATGCGCAAGCTGCTAATTTTCATTTAAATTTTACCAACAGGCTGTAATGG 727
Qy 731 CTGATGAATGGAATGCAATATACCTCTTCAAAATTTGAACTTAATGCTGGAACATCAG 790
Db 728 CTGATGAATGGAATGCAATATACCTCTTCAAAATTTGAACTTAATGCTGGAACATCAG 787
Qy 791 ATGACTATTAATTAATTTTAAAGAAATATACCTTAATATAGTAATGATGCAATA 850
Db 788 ATGACTATTAATTAATTTTAAAGAAATATACCTTAATATAGTAATGATGCAATA 847
Qy 851 CCTATAGAACCACTAATAAAATCTTAGAGAGAACCAATATGAAATGAGATATTTA 910
Db 848 CCTATAGAACCACTAATAAAATCTTAGAGAGAACCAATATGAGATGAGATATTTA 907
Qy 911 ATGACTATGGAATATATGACCATTTACTGATTTAGATACCATCTGCAATTTTCTTAT 970
Db 908 ATGATTTATCGAAGATATATGACTATTACTGATTTAGATACCATCTGCAATTTTCTTAT 967
Qy 971 ATGATTTAATAAAGATATAGATTCATATGAGAGAAATGAAATGAAGCAATTAAGATG 1030
Db 968 ATGATTTAATAAAGATATCAAGATTCATATGAGAGAAATG-----GTGGCATTAATACTG 1021
Qy 1031 AACTCACAGAGAAATTTATCAACTGAAATTAATTTGATCGTCTCTCAACTTAGAG 1090
Db 1022 AACTTACAGAGAAATTTATCAACTGAAATTAATTTGATCGTCTCTCAACTTAGAA 1081
Qy 1091 TTCAACCCCAATCTAGCTAGATGGAATATATTAACAGTGCAGAGTTTAAATTTAT 1150
Db 1082 TTCAACCCCAATCTCGCTAATATGGAATTAATTTAACAAGTGCAGAGTTTAAATTTAT 1141
Qy 1151 CATTTTGAACAATTTATTTTATTAAGAAAAACAATTCGGGAATCGTTAGTTG 1210
Db 1142 CATTTTGAAGAACTTATATTTATTAACAAAAAGAAACGACGGGAATCGTTAGTTG 1201
Qy 1211 GTATTTCTAATCGTATGACCACTTACTTATAGCAATACTATTAAGAACTTTATAGAG 1270
Db 1202 GTATTTGGAATCGTATATGATCTTATATGCTACGACAGGAATGTAATATATAGAG 1261
Qy 1271 AAAGAACAGGTTCCACCAACAAACAAATAGACCAATTTGAATCTTATTAAGTTCAA 1330
Db 1262 AAAGAACAGGTTCCACCAACAAACAAATTTATATCCATTTGAATCTTATTAAGTTCAA 1321
Qy 1331 TTGTAACTGATAGCAATCACTCTGTTTCCCAATTC---AACCAACTTTATTAATTA 1387
Db 1322 TTGTAACTGATAGCAATCACTCTGTTTCCCAATTTCTTAAACATTAACCTTTCAATTA 1381
Qy 1388 ATCAAAATGAACCTTATTTAAATGCGCTATCAACACACTCAAAATATTCAGAGAG 1447
Db 1382 ATCAAAATGAACCTTATTTAAATTAATTCACCTAGTAATTAATTAATTCAGAGAG 1441
Qy 1448 GGTCTTTATCTAATATGCAAAACACAACTTTTTCAAATTTCTAGAAAAAAGACTGCA 1507
Db 1442 GGAATTTATCTAATATGTAATAAACAACGATTTTCAATTTCTGTAAAAAAGACTGTA 1501
Qy 1508 ATCTAGTTATGATTCAGAGTTGTTCAACAACTTTAATTAATTAATTAATTTATCC 1567
Db 1502 AACCAATTTATTAATTCAAATGTTTACCAAGCTAATTAATTAATTAATTTATCC 1561
Qy 1568 ATTTTTCATTTTACTATTCCTATGATGATGATGATGATGATGATGATGATGATGATGAT 1627
Db 1562 AGTTTTCATTTTATTAATTCCTATTAATAATTTGATGATGATGATGATGATGATGATGATGAT 1621

Qy 1628 GTGATTTAGATGAGACACACAGTATGTTGATATGATATTAATGCAATATCAGATTAATA 1687
Db 1622 GTGATTTAGATGAGACACACAGTATGTTGATATTAATGCAATATCAGATTAATAATA 1681
Qy 1688 TTCAATGATCCCGAGCAATCAAGGTAAACAATCTGTGATCAAACTCTAAGGTAATTAAG 1747
Db 1682 TTCAATGATCCCGAGCAATCAAGGTAAACAATCTGTGATCAAACTCTAAGGTAATTAAG 1741
Qy 1748 GACCTGTATACAGAGAGAACTTGTTTATTTTACAAAGTCAAGGCGTTTAGAAATTA 1807
Db 1742 GACCTGTATACAGAGAGAACTTGTTTATTTTACAAAGTCAAGGCGTTTAGAGATTA 1801
Qy 1808 CATGTGAATCTCTAATTTCTACACAACTTTATTTCAATTTAGACTTGAATATGCAAAATG 1867
Db 1802 CATGTGAATCTCTAATTTCTACACAACTTTATTTCAATTTAGACTTGAATATGCAAAATG 1861
Qy 1868 GTGCTGGAATATCTCTCTAATATATCTCTTCAATACAGAGAGTAATGAAATACAC 1927
Db 1862 GTGCTGGAATATCTCTCTAATATATCTCTTCAATACAGAGAGTAATGAAATACAC 1921
Qy 1928 CTCAACGACTCAACACACTTTTCTGTACAAATTTAATTAATTTTCAATACGAGATT 1987
Db 1922 CTCAACGACTCAACACACTTTTCTGTACAAATTTAATTAATTTTCAATACGAGATT 1981
Qy 1988 TTGGGTAATTTCCAAATTTCCAGTACAGTAACTTACCTTTAATCGAAACATACCAATTTA 2047
Db 1982 TTGGGTAATTTCCAAATTTCCAGTACAGTAACTTACCTTTAATCGAAACATACCAATTTA 2041
Qy 2048 TATTTATCGGACAGATATCAAAATTCATTTTATCATTTGAATTAATTTATAC 2107
Db 2042 TATTTATCGGACAGATATCAAAATTTCAATTTTATCATTTGAATTAATTTATAC 2101
Qy 2108 CAATTTACTCTCTATGACCAACAAATAGAGAAAAACAAATTTAGAAATTAATCCAAACA 2167
Db 2102 CAATTTACTCTCTATGACCAACAAATAGAGAAAAACAAATTTAGAAATTAATCCAAACA 2161
Qy 2168 AATTAATTAATTTTTCACCAATATATCAAAACACTTTGA 2208
Db 2162 AATTAATTAATTTTTCACCAATATATCAAAACACTTTGA 2202

RESULT 3
ADR89398
ID ADR89398 standard; cDNA; 2085 BP.
XX
AC ADR89398;
XX
DT 18-NOV-2004 (first entry)
XX
DE AXMI-007 alternative start site coding sequence.
XX
KW ss; gene; delta-endotoxin; delta-endotoxin associate polypeptide;
KW expression cassette; transformation; transgenic; plant; bacteria;
KW lepidoptera; coleoptera; pest; pesticide; resistance;
KW pesticial activity.
XX
OS Bacillus thuringiensis.
XX
FH Key location/Qualifiers
FT 1..2085
FT CDS /*tag= a
FT /product= "Alternative AXMI-007"
XX
PD MO2004074462-A2.
XX
PD 02-SEP-2004.
XX
PF 20-FEB-2004; 2004MO-US005829.
XX
PR 20-FEB-2003; 2003US-0448632P.
PR 20-FEB-2003; 2003US-0448633P.
PR 20-FEB-2003; 2003US-0448797P.

20-FEB-2003; 2003US-0448806P.
 20-FEB-2003; 2003US-0448810P.
 20-FEB-2003; 2003US-0448812P.
 19-FEB-2004; 2004US-00781979.
 19-FEB-2004; 2004US-00782020.
 19-FEB-2004; 2004US-00782096.
 19-FEB-2004; 2004US-00782141.
 19-FEB-2004; 2004US-00782570.
 19-FEB-2004; 2004US-00783417.
 (ATHE-) ATHENIX CORP.

Carozzi N, Hargis T, Koziel MG, Duck NB, Carr B;
 WPI; 2004-635574/61.
 P-PSDB; ADR89399.

New isolated delta-endotoxin and delta-endotoxin-associated nucleic acids
 and polypeptides, useful for killing lepidopteran or coleopteran pests or
 for producing organisms with pesticide resistance.

Claim 1; SEQ ID NO 10; 178bp; English.

This sequence encodes an isolated delta-endotoxin. Some of the delta-
 endotoxin coding sequences of the invention have alternative start
 codons, producing more than one protein from a single open reading frame.
 The nucleic acid sequences of the invention are useful in DNA constructs
 or expression cassettes for transformation and expression in plants and
 bacteria. The nucleic acids and corresponding polypeptides are useful for
 killing lepidopteran or coleopteran pests. Compositions containing the
 delta-endotoxins of the invention, and methods for their production, are
 useful for the production of organisms with pesticide resistance,
 specifically bacteria and plants. These organisms are useful for
 generating altered or improved delta-endotoxin or delta-endotoxin-
 associated proteins that have pesticidal activity, or for detecting the
 presence of delta-endotoxin or delta-endotoxin-associated proteins or
 nucleic acids in products or organisms.

Sequence 2085 BP; 794 A; 338 C; 302 G; 651 T; 0 U; 0 Other;

Query Match 81.3%; Score 1794.4; DB 13; Length 2085;
 Best Local Similarity 92.7%; Pred. No. 1.9e-302;
 Matches 1908; Conservative 0; Mismatches 141; Indels 9; Gaps 2;

154 ATGTGTCAGGGAATACCAATATGATGATATTTGAGACATTTGCTAGTCTGATCA 213
 1 ATGTGTCAAGGGAATACCAATATGATGATATTTGAGACATTTGCTAGTCTGATCA 60
 214 ATTTGTCAGGTAATGACAGTACTATTTGATCCGGTACTCTGTAGCCGTAATAGTGG 273
 61 ATTTGTCAGGTAATGACAGTACTATTTGATCCGGTACTCTGTAGCCGTAATAGTGG 120
 274 CTGACTTCTATATCCGGAACGATAGGAATATATGATGATATATCTTTGGTACC 333
 121 CTGACTTCTATATCCGGAACGATAGGAATATATGATGATATATCTTTGGTACC 180
 334 CTAACTACTGTTTGGCCGCGGGAACAAAGCAAGTATGACACAAATTTATT 393
 181 CTAACTACTGTTTGGCCGCGGGAACAAAGCAAGTATGACACAAATTTATT 240
 394 AAAATGGAGAAATTTTGTGATACCGTTAATAGAAAGCATTAACAGTTAAAGTTA 453
 241 AAAATGGAGAAATTTTGTGATACCGTTAATAGAAAGCATTAACAGTTAAAGTTA 300
 454 CAACTTTAGAGATTTAGACAAATATTTCAAAAGCTATATATAGCATTTAGATTTGG 513
 301 CAACTTTAGAGATTTAGACAAATATTTCAAAAGCTATATATAGCATTTAGATTTGG 360
 514 AGAAATTTAAAGATCAAGCTCTGATTAACACCATCATCTAGCATTTAAAGAGT 573
 361 AGAAATTTAAAGATCAAGCTCTGATTAACACCATCATCTAGCATTTAAAGAGT 420
 574 GCCTTGACTTTAAATAGATTTGAGAAATGTTCAAAATGATTTTATTCGAAATATCT 633

421 GCCTTGACTTTAAATAGATTTGAGAAATGTTCAAAATGATTTTATTCGAAATATCT 480
 634 GGTTCACACTTGAACCTTATTAACCGCTATTAAGTCTATTTATGCGCAAGCTGAT 693
 481 GGTTCACACTTGAACCTTATTAACCGCTATTAAGTCTATTTATGCGCAAGCTGAT 540
 694 TTTCAATTTAAATTTATTAACCAAGGTCGTAATTTGCTGATGATGATGATGATGAT 753
 541 TTTCAATTTAAATTTATTAACCAAGGTCGTAATTTGCTGATGATGATGATGATGAT 600
 754 CATCTTGAACAAATTTGAACCTTATGCTGGAACATGATGATGATGATGATGATGAT 813
 601 CATCTTGAACAAATTTGAACCTTATGCTGGAACATGATGATGATGATGATGATGAT 660
 814 GAAATATATACCTTAATATGATGATGATGATGATGATGATGATGATGATGATGAT 873
 661 GAAATATATACCTTAATATGATGATGATGATGATGATGATGATGATGATGATGAT 720
 874 CTTAGAGACGAACCAATATGATGATGATGATGATGATGATGATGATGATGATGAT 933
 721 CTTAGAGACGAACCTTAATATGATGATGATGATGATGATGATGATGATGATGATGAT 780
 934 ATTACTGATTTAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 993
 781 ATTACTGATTTAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 840
 994 TCAATAGAGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1053
 841 TCAATAGAGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 894
 1054 ACTGAAATATATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1113
 895 ACTGAAATATATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 954
 1114 GAAATATATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1173
 955 GAAATATATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1014
 1174 TATACGAATATACAAATTTGGGGAATGCTTATGATGATGATGATGATGATGAT 1233
 1015 TATACGAATATATGAAACCTAGGGAATGCTTATGATGATGATGATGATGATGAT 1074
 1234 ACTTATGACATATATATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1293
 1075 ACTTATGACATATATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1134
 1294 AAAACATATGACATATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1353
 1135 AAAACATATATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1194
 1354 CCGTTTCCCTATTC--AACCACTTTATATATGATGATGATGATGATGATGAT 1410
 1195 CCGTTTCCCTATTC--AACCACTTTATATGATGATGATGATGATGATGATGAT 1254
 1411 GGCCTATATACACACCTGAAATATATGATGATGATGATGATGATGATGATGAT 1470
 1255 AATTCACCTATATATATGATGATGATGATGATGATGATGATGATGATGATGAT 1314
 1471 ACAAATTTTTCATTTCTAGAAAAAAGCTGCAATCTAGTATTTGATGATGATGAT 1530
 1315 ACAAATTTTTCATTTCTAGAAAAAAGCTGCAATCTAGTATTTGATGATGATGAT 1374
 1375 TTACCAAGCTATATATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1434
 1591 TATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1650
 1435 TATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1494
 1651 AGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1710

```
Db 1495 AGTGTATAGAAATATGCAATATCAGATATAAATTAATTACATGATGCCAGCATCAA 1554
Oy 1711 GGTAACAATCTTGATGACAACTCTAAGTAATTGAAGACCTGGTCATACAGAGGAAAC 1770
Db 1555 GGTAAACGCTCTGATACCAACTCTAAGTAATTGAAGACCTGGTCATACAGAGGAAAC 1614
Oy 1771 TTGGTTATTTACAAAGTCAAGGGCGTTTGAATTAATGATGAAACCTCTAATTCTACA 1830
Db 1615 TTGGTTATTTACAAAGTCAAGGGCGTTTGAATTAATGATGAAACCTCTAATTCTACA 1674
Oy 1831 CAATCTTATTTCAATGAGCTTGCATATGCTCAAAATGGTGTGGAATTAATCTCTTCAAT 1890
Db 1675 CAATCTTATTTCAATGAGCTTGCATATGCTCAAAATGGTGTGGAATTAATCTCTTCAAT 1734
Oy 1891 ATATCTCTTACATACCAAGGATTAAGAAATACCACTCAACGATCAACAAACATTTT 1950
Db 1735 ATATCTCTTACATACCAAGGATTAAGAAATACCACTCAACGATCAACAAACATTTT 1794
Oy 1951 TCTGTTACAAATTTAATTAATTTACAAATACGAGATTTTGGGTATTTCCAAATTTCCAA 2010
Db 1795 TCTGTTACAAATTTAATTAATTTACAAATACGAGATTTTGGGTATTTCCAAATTTCCAA 1854
Oy 2011 ACAGTACATTAACCTTTAAATCGAAAACATACATTTATTTAAATCGTACAGATGATCA 2070
Db 1855 ACAGTACATTAACCTTTAAATCGAAAACATACATTTATTTAAATCGTACAGATGATCA 1914
Oy 2071 AATTCATTTTAACTGATTAATTAATTTGAATTTATTAATTAATTAATTAATTAATTAAT 2130
Db 1915 AATTCATTTTAACTGATTAATTAATTTGAATTTATTAATTAATTAATTAATTAATTAAT 1974
Oy 2131 AATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2190
Db 1975 AATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2034
Oy 2191 CATACAAAACACCTTTGA 2208
Db 2035 CATACAAAACACCTTTGA 2052

RESULT 4
ID AAN93054 standard; DNA; 3940 BP.
XX
AC AAN93054;
XX
DT 27-AUG-2003 (revised)
DT 25-MAR-2003 (revised)
DT 30-MAY-1990 (first entry)
XX
DB Delta-endotoxin crystal protein gene.
XX
KW Delta-endotoxin; crystal protein; insecticide; ss; pcc130;
KW biological control agent.
XX
OS Bacillus thuringiensis; israeliensis.
XX
FH Key Location/Qualifiers
FH RBS 879..884
FH /tag= b
FH /label= Shine-Delgarno sequence
FH CDS 891
FT /tag= a
FT /product= "delta-endotoxin"
XX
XX EP296870-A.
XX
XX 28-DEC-1988.
XX
XX 24-JUN-1988; 88EP-00305772.
XX
XX 26-JUN-1987; 87US-00067653.
XX
XX (DUPO ) DU PONT DE NEMOURS & CO E I.
```

```
XX
FI Ellar DJ, Ward ES;
XX
XX WPI; 1989-001322/01.
DR P-PSDB; AAP94035.
XX
XX DNA fragment encoding insecticidal protein - obtd. from Bacillus
FT thuringiensis sub species israeliensis, and used in microorganisms and
PT plant cells.
XX
PS Disclosure; Fig 5; 26pp; English.
XX
CC The nucleotide sequence is an insert in plasmid pcc130. The delta
CC endotoxin protein is insecticidal and can be used to control insect pests
CC esp. mosquitoes. See also AAN93059. (Updated on 25-MAR-2003 to correct PA
CC field.) (Updated on 27-AUG-2003 to correct OS field.)
XX
SQ Sequence 3940 BP; 1457 A; 603 C; 659 G; 1221 T; 0 U; 0 Other;

Query Match 15.7%; Score 346.6; DB 1; Length 3940;
Best Local Similarity 53.2%; Pred. No. 6e-51;
Matches 1176; Conservative 0; Mismatches 934; Indels 99; Gaps 17;

Oy 4 AATCAAAATACCATTAATACGAATATGAATTTATGATTCGATTAACCTTATTTT 63
Db 888 AATATGAATCCTTATCAAAATTAATAATGAATGAATGAATTAATTAATTAATTAATTAAT 947
Oy 64 CCGAAGCAAGAAACAGTATGATTTAGATTAACCTTACCAAAATATCCAAATCAACATTA 123
Db 948 TTAATATATCTAATATTAATTAATTAACAGATATCCAAATGAATATGTCAAAATTAATTA 1007
Oy 124 CAAAACCAAAATTAACAAAGAGTGGCTCAATATGTTGTCAAGGAATATACATATGTTGAT 183
Db 1008 CAAAATCAAAATTAATAAGATTTGGCTCAATATGTTGTCAACAGATACGAGTATGTGA 1067
Oy 184 AATTGGAGCATTTGCTAGTGTGATACAAATTCGTCAGTTAGTGGAGGACTATTGTA 243
Db 1068 GATTTGAACCTTTATATGAT-----AGTGTAACTCAGTGTCTTATTAATTAATTAAT 1118
Oy 244 TCCGTTACTCTGTTAGCCGGTATAGTGGGCTCACTTATATCCGACCGATAGATATA 303
Db 1119 GTTGGACCGTACTGACTGGT-----TTCCGGTTCCACAACACC 1157
Oy 304 ATAGGTCTATTAATATATCTTTTGGTACCTTAATCACTGTCTTTTGGCCCGGGAGAA 363
Db 1158 TTAGGACTTCTTAATATAGGTTTGGTATTAATTAATCAATGTTCTTTTCCAGCCCAAGAC 1217
Oy 364 CAAGACAAACAGATATGACACAAATTTATTAATGAAGGAATTTTGTGATACACCG 423
Db 1218 CAATCTAACAC--ATGGAGTACTTTATTAACCAAACTAATAATTAATTAATAAAGAA 1274
Oy 424 TTAAACA-GAAGCATTAATAACAGCTTAAGTTACAACCTTTAGAGATTTAGACAAATATT 482
Db 1275 ATAGCATCAACATATATATAGTAATGCTAATTAATTTTAACAGTGGTTTAAGTTATATC 1334
Oy 483 ACAAGCTATTAATACAGCATTAATGATGATTTGAGAGAAATTAATAAGACTACAGCTCTCG 542
Db 1335 AGCACTTATATATATACCTTTAAACATGGAGATATATCCAAACCAAAATATATCTCAG 1394
Oy 543 ATTACACCATCATACAGATTAACAAGAGTGCCTTACCTTAAATATAGATTGGA 602
Db 1395 GATGTAAAGA-----CACAAATCCAGCTAGTTCAATTAACCAATTTCAAAATGTCCACA 1450
Oy 603 TGTTCACAATGATTTTATTCGAGAAATACCTGGTTTCCAACTTGAAACTTAATAACGCT 662
Db 1451 GCTTGAATCTCTGTCTCTTAATCTTAATGATGCTG-----GATTAATTAATCAATCT 1504
Oy 663 ATTACTACTAATTTATGCGAAGCTGCTAATTTTCAATTTAATTTATTAACAAGGTGC 722
Db 1505 AGTATTAATCTAGTATGACAAAGCAGCAAACTTAATCTGATGATTAATTAATCAAGCGT 1564
Oy 723 TGAATTGGCTGATGAATGAATGAGATGAGATTAATCTTCAATTAATTAAGTATGCTGG 782
```

```

Db 1565 CAAATTTGAGGATTTAAAAACAATCGACATTCGATTTATTTAGACCTTTGCC--- 1621
Oy 783 AACATCAGATGACTATTTAACTTTTAAAGAAATATACCTAATATATGATACCTTTG 842
Db 1622 AACGCAATTTGATTTATTCAGTATTTGCTAAGCTATAGAAATTTACATTTATTTG 1681
Oy 843 TCGAAATACCTATAGACAGACTAAAAATCTTAGACGACAAACCAATATGAAATGAG 902
Db 1682 TGTAAACAATTATAAAAAGATTTAAATTTAATTAAGACGCTGATAGTATCTTGA 1741
Oy 903 TATATTTATGACTATCGAAGATATATGACCTATCTGATTTATGATACCTCTCAAT 962
Db 1742 TCGAAATTTAACTCGAAACACATA-----CAATCGATCGAAACAAATATCACTAC 1795
Oy 963 TTCTTTATATGATATAAAAAGATATAGATTTCAATAGAGGATATAGAAAGGAT 1022
Db 1796 TGTATTTAATGTTGTTGCACTCTTCTTATATATGATGATGATTAATATCCATAGGT 1855
Oy 1023 TAAGATGACTCAAGAGAAATTTATACAAGTAAATTTTGAATCTGCTCTCA 1082
Db 1856 CCAATCTGACTTACTCGAAGAAATTTAT--CAGTACTTAACTTGAAGAAAGCCCTAT 1913
Oy 1083 ACTTGAAGTTAAACCAATCTAGCTAGAGATGATATATTTAACAGTGAAGTTTAA 1142
Db 1914 AAATATTTATGACTTTCATATCAAGAGATTCATTAACGTAAGAC--GCATTTATTTA 1971
Oy 1143 ATATTTTCAATTTTGAACAATTTATTTTATACAGAAATAGAAATTTGCGGAATCG 1202
Db 1972 CTGGCTGATCTTTGAA-----TTTATATGAAAAAGGCAAACTACTCTTAATTAAT 2026
Oy 1203 TTTAGTTGATTTTCTAAATCGTATGACCTTACTTATAGCACTATATACGAAACTTT 1262
Db 2027 TTTACACGACATTTATATATGTTTCAATTAACACTGATATATATCCAAAAATCTAG 2086
Oy 1263 ATATGAGAAAGAACAGGTTCAACCAACAAAAATTTAGACCAATTTGAATCTTATTA 1322
Db 2087 TGTATTTGAAATCACAATGTATATGATTAATTAATCTCTGTTGGCAAAATAT 2146
Oy 1323 AGTTTCATTTGTAATGATAGACATCACTCTCTTTCCCTATTTCAACCACTTTAT 1382
Db 2147 TTTATTTTATTTTAAATGTCATAGCTTATATATTAATATCTTAATGATTAATTA 2206
Oy 1383 AATTATCAAAATTTGAATTTTATTAATGCTCTATCAACACACTCAATATTTAGC 1442
Db 2207 TATTTAGTAAATGATTTTATTAATGATGATCTAGACTTTTGGAGAAAGACTTAC 2266
Oy 1443 AGAGAGGCTCT-----TATCTAATTTCAAAAACAATTTTTCATTTCTAGAAA 1496
Db 2267 AGCAGGATCTGGGCAAAATTAATCTTATGATGTAAATAAATATTTTCGGGTTACCAATCT 2326
Oy 1497 AAAAGACTGCAATCTAGTATTTGATCCAGGTTTCAACCAACTTTAATTAATTAATGTC 1556
Db 2327 TAAACGAAGAGAAATCAAGAAACCTTACCTTTTCCAAATATGATTAATGATGTC 2386
Oy 1557 TATTTTATCCATTTTCTATTTACTTATTTCTATGATGATTTGATTAACAGTCAAT 1616
Db 2387 TATTTTATC-----ATTATTAATAAGCTTATGATTCCTGCAACATAT 2429
Oy 1617 ATTAGATACAGGTGTA--TTAGATGACACACAGTATGTTGATATATATATGCAATTA 1674
Db 2430 AAAACTCAAGTGTATACGTTGCTTGACACACTCTAGGTTGATCTTAATAATTAATCAAT 2489
Oy 1675 TCAGTAAATTAATTAACATATGCCAATCAAGATTAACAACTTGATTAACAATCT 1734
Db 2490 TATACACTTTTAACCTAACCAATTCAGCTGTAAAGGAATTCCTGGGACTGCTCTCT 2549
Oy 1735 AAGTAAATTTGAAGACTGCTGCTATACAGAGGAACTTGGTTATTTTCAAAAGTCAAGGG 1794
Db 2550 AAGTGTGTTCAAGGACTGCTGCTATACAGAGGGAATTTAATTTGATTTCAAGATCA---- 2605
Oy 1795 CGTTTGAATTTACATGTGAAATCTCTAATTTTACACATCTTATTTTCAAGCTTGA 1854
Db 2606 -TTTCAAAATTAACATGTCAACACTCAAAATTTTCAACATCTGATTTTATTAAGAAATGCT 2663

```

```

Oy 1855 TATGCTACAAATGCTGTGAAATATCTCTCTAATATATCTTACAAATACAGAGTA 1914
Db 2664 TATGCTTAAATGAGAGGCAATATCTGAGCTGTTATTAATCTTATGATCCAGGG-- 2721
Oy 1915 ATAGAAATACCACTCAAGACTCAACAAACTTTTCTGTGTACAAATTAATTAATTTA 1974
Db 2722 -TAGCAAGACTGGGATATGACACTCAACCCCACTTTTCTGGTACAGATTAATGAAATTTA 2780
Oy 1975 CAATTCGAGATTTTGGGTATTTTCCAAATTTCCAACTACAGTAACTTACCTTAATGCA 2034
Db 2781 AAATATTAAGATTTTCACTACTTAAATTTTCTAACGAGGAAATTTGCTCAAAATCAA 2840
Oy 2035 AACATACATTTATATTTATATGTCAGATGAT---CAAAATTTATTTATCAATGAT 2091
Db 2841 AACATATCTCTTGCTTTATATGCTTGGATGATATATACAAACACACAGTACTTATTTGAT 2900
Oy 2092 AAAATGAAATTTATACCAATTTACTTCTATGACCAAAATAGAGAAACAAAAATTTA 2151
Db 2901 AAAATGAAATTTTCTGCAATTTACTGTTCTATTAAGAGAGATAGAGAAACAAAAATTTA 2960
Oy 2152 GAACTATCCAAACAAAAATTAATATCATTTTTCACAAATCATACAAAA 2200
Db 2961 GAAACAGTACAAACAAATTAATTAATCATTTTATGCAAAATCCTATTAATAA 3009

RESULT 5
AAN93059
ID AAN93059 standard; DNA; 4571 BP.
XX
AC AAN93059;
XX
DT 27-AUG-2003 (revised)
DT 25-MAR-2003 (revised)
DT 30-MAY-1990 (first entry)
XX
DE Delta-endotoxin crystal protein gene.
XX
KM Delta-endotoxin; crystal protein; insecticide; pCH130; ss;
biological control agent.
XX
OS Bacillus thuringiensis; israeliensis.
XX
FH Key location/Qualifiers
FT RBS 879..884
FT /tag= b
FT /label= Shine-Dalgarno sequence.
FT misc_feature 891..4430
FT /*tag= a
FT /product= "delta-endotoxin crystal protein"
XX
FN EP296870-A.
XX
PD 28-DEC-1988.
XX
PF 24-JUN-1988; 88EP-00305772.
XX
PR 26-JUN-1987; 87US-00067653.
XX
PA (DUPO ) DU PONT DE NEMOURS & CO B I.
XX
PI Eljar DJ, Ward ES;
XX
WPI, 1989-001322/01.
XX
DR P-PSDB; AAP93715.
XX
PT DNA fragment encoding insecticidal protein - obcd, from Bacillus
thuringiensis sub species israeliensis, and used in microorganisms and
plant cells.
XX
PS Disclosure; Fig 7; 26pp; English.
XX
CC The sequence encodes the 130 kDa delta-endotoxin gene from B.

```


cc thuringiensis subsp. israeliensis and is a 4.46 kb insert of pCH30. The
cc delta endotoxin protein is insecticidal and can be used to control insect
cc pests esp. mosquitoes. See also AN93054. (Updated on 25-MAR-2003 to
cc correct PA field.) (updated on 27-AUG-2003 to correct OS field.)

xx Sequence 4571 BP; 1678 A; 685 C; 817 G; 1391 T; 0 U; 0 Other;

Query Match 15.7%; Score 346.6; DB 1; Length 4571;
Best Local Similarity 53.2%; Pred. No. 6e-51;
Matches 1176; Conservative 0; Mismatches 934; Indels 99; Gaps 17;

```
QY 4 AATCAAAATTAAGATAATAGCAATATAGTAATTAATGATGCGATACCTGACTTATTTT 63
DB 888 AATATGAAATCTTATCAAAATTAATTAATGAAATTAATGAAATTAATGCTTCAAAAAA 947
QY 64 CCGAAGAGAAACAGTAATGATTTAGATACCTTTACACAAATTAATCCAAATCAACATTA 123
DB 948 TTAATATATCTAATTAATTAATCAAGATATCCAATAGAAAATATGCCAAAATTAATTA 1007
QY 124 CAAAACAAATTTACAAAAGAGTGGCTCAATATGTCTCAAGGAAATACAAATATGTGTAT 183
DB 1008 CAAAGTCAAAATTAATTAAGATTTGGCTCAATATGTCTCAAGAAATCAAGAGATGTGGA 1067
QY 184 AATTGAGACATTTGCTAGTGTGATACAAATGCTGAGTTAGTCAAGTACTATTGTA 243
DB 1068 GATTTGAACTTTTATGAT-----AGTGTGAACTCAAGTGCCTATATCTATTGTA 1118
QY 244 TCCGTAATCTGTTAGCCGATATAGTGGGCTCACTTATATCCGAGCCGATAGATA 303
DB 1119 GTTGGAGACGTAAGTACTGT-----TTGGGTTCAACAACCC 1157
QY 304 ATAGTGTCTAATTAATATCTTTTGGTACCTTAATCACTGTCTTTGGCCCGGGAGAA 363
DB 1158 TTAGAGACTGTCTTAATATGTTTGGTATCAATTAATACAGTTCCTTTTCCAGCCCAAGAC 1217
QY 364 CAAGCAAAAGATATGGAACAATTAATTAATTAATGGAATTTTGTGATATACCG 423
DB 1218 CAATCTAACAC---ATGAGTGAATCTTAATACAACTAAATAATTAATTAATAAAGAA 1274
QY 424 TTAACA-GAAAGATTAACAGCTAAAGTTACAACTTAGAGATTTAGACAAATAT 482
DB 1275 ATAGATTAACATATATATAGTAATGCTAATTAATTAATTAAGAGTGCCTTAATGTTATC 1334
QY 483 ACAAACTATATATACAGATTAATGATGAGAAATTAATAAAGACTACAGCTCTGCG 542
DB 1335 AGCACTTATCAATATCACTTAATAACATGGAAGATATCCAAACCCCAAAATATACAG 1394
QY 543 ATTAACAATCATCATGAGATTAACAAGCTGCCCTGATCTTAATAATACATTTGAGAA 602
DB 1395 GATGTAAAGAA---CACAAATCCAGCTAGTTCATTAACATTTTCAAAATGTCATTTCCAGA 1450
QY 603 TGTTCAATGATTTTATTCGAAATATACCTGTTTCCAACTTGAATTTAAACGCT 662
DB 1451 GCTGTAAATCTGTGCTCTCCATTCCTAGTATTCG-----GATTAATTAATACATCT 1504
QY 663 ATTACTACCTATTTATGCGCAAGCTGCTAATTTTCATTTAAATTTATTAACAAGCTGC 722
DB 1505 AGTATTAATCTATGATGCAAGAGCAAGCAACTTACATCTGACTGATTAATTAATCAAGCGT 1564
QY 723 TGAATGCTGATGAATGGAATGAGATATACATCTTCACAATTAATGAACCTAATGCTGG 782
DB 1565 CAATTTAAGCGTATTTAAATAAACAATGCAATTCGATTTATTAAGCCCTTTGCG--- 1621
QY 783 AACATCAGATGATTTAATTAACCTTTTAAAGAAATATACCTAAATTAATTAATGAT 842
DB 1622 AACGCAATGATTTATTAATCAAGTATTAATTAAGCTATTAAGATTTACATTAATTTG 1681
QY 843 TGCAAAATCTATTAAGACAGATTAATAAATCTTAAGAGCAACCAATTAATGAATGAG 902
DB 1682 TGTAAACAATTAATAAAGATTAATTAATTAATAAGACGCGCTGATTAATCTTTGA 1741
QY 903 TATATTTAATGATATGCAAGATATATGACATTAATCTGATTAATGATACATCTCTCAAT 962
```

```
DB 1742 TCGAAATATAAATCTGGAACACAT-----CAATAGCTATCGAACAATAATGACTCTGC 1795
QY 963 TTCTTTATATGATATTAATAAAGATATAGAGATTCATATAGAGATATAGAGTAAAGCAT 1022
DB 1796 TGTATTAATGATGTTTGCACCTTTCTTAATTAATGATGTAATGTAATATCCAAATAGGTGT 1855
QY 1023 TAGAATGAATCTCAAGAGAAATTTTATACACTGAAATTAATTTTGAATGCTCTTCTCA 1082
DB 1856 CCAATCTGAATCTTACTGGAATAATTTAT--CAGGTACTTAATCTTGAAGAAAGCCCTAT 1913
QY 1083 ACTAGAGTTCAACCCATCTAGCTAGATGAATATTAATTAACAGTGCAGTTTAA 1142
DB 1914 AATATTAATGACTTTCAATTAAGAGATTTCACTTACAGTGAAC--GATTTATTTA 1971
QY 1143 ATTAATTTCAATTTTATAGAACAAATTTATTTTATACGAAATAACAAATTTGGGAAATCG 1202
DB 1972 CTGGCTTGAATCTTTGAA-----TTTTTATGAAGAAAGCGCAACCTACTCTTAATTAAT 2026
QY 1203 TTTAGTGTATTTCTAATGCTATGACCTATCTTAATAGCAATCTTAATCTGAACCTTT 1262
DB 2027 TTTCAACGACATTAATATATATGTTTCATTAACACTGATATATATATCCAAAAATCTAG 2086
QY 1263 ATATGAGAAAGAACAGTTTCAACCAACAACAAAAACAATTAAGACATTTGATCTTAAT 1332
DB 2087 TGTTTTGAAATCAATGATATGATTAATTAATTAATCTTGTGTTGGCAAAATAT 2146
QY 1323 AGTTCAATTTGATCTGATAGACATCACTCTGTTCCCTATTCACACACTTAT 1382
DB 2147 TTAATTTTATTAATGATCAATAGCTTAGATATTAATTAATCTAATATGATTAATTA 2206
QY 1383 AATTAATCAATTTGAATTTATTAATGCTCATCTTAACAACACTCAATTTTCAGC 1442
DB 2207 TATTAAGTAATATGATTTTATTAATTAATGATTAATGATTAATGATTAATGATTA 2266
QY 1443 AGGAGGCT-----TTATCTAATTAATCAAAACAACTTTTTCATTTCTCTAGAA 1496
DB 2267 AGCAGATCTGGGCAATTAATCTTAATGATTAATTAATTAATTTTGGGTTTCCAAATCT 2326
QY 1497 AAAAGCTGCAATCTAGTATTAATGATCCAGGTGTTCCAAACTTATTAATTAATTAATGCA 1556
DB 2327 TAAACGAAGAGATCAAGAAACCTTACCTTTTCCAAATATGATTAATCTATGATCA 2386
QY 1557 TATTTATCCCATTTTTCATTAATTTACTTAATCTTATGATGATGATTAATCACTCAAT 1616
DB 2387 TATTTATTC-----ATTTATTAAGTCTTATGATCCCTGCAACATAT 2429
QY 1617 ATTAGATACAGTGA--TTAGAGTGAACAACAGTATGTTGATTAATTAATGACATA 1674
DB 2430 AAAATCTAAGTATATGATTTGCTTGAACACCTTAATGATTTCTTAATAATTAATTA 2489
QY 1675 TCAGATTAATAATTAATCAATGATCCAGCAATCAAGGTAACAATCTTGAATCAACCTCT 1734
DB 2490 TATACATATTAATTAATCCAAATTCAGCTGTAAAGCAATTAATCACTTGGGAGCTTCT 2549
QY 1735 AAGTATTTGAAGACCTGCTCATACAGAGAACTTGGTTTATTTTAAAGTCAAGG 1794
DB 2550 AAGGTGTTCAAGACCTGCTCATACAGAGGGAATTTAATGATTTCAAGATCA---- 2605
QY 1795 CGTTAGAAATTAATGATGAACTCTTAATTTCTACAACTTATTTTCAATTAATGATGAT 1854
DB 2606 --TTTCAAAATTAATATCACTCAACTCAAAATTTTCAACATGATTTTATTAAGATTCGT 2663
QY 1855 TATGCTAACAATGCTGTAATAATCTCTTAATTAATCTTAAATTAACAGAGTA 1914
DB 2664 TATGCTTAATTAATGAAGCAAGCAATTAATCTGAGCTGTTAATCTTAATCTTAATCCAGGG-- 2721
QY 1915 ATAGGAATACACCTCAACGATCAACAACATTTTCTGATCAATTAATTAATTAATTA 1974
DB 2722 -TAGCAAACTGGGATATGCACTCAACCCCACTTTTCTGATCAAGTATTAACAAATTA 2780
QY 1975 CAATAGGAGATTTTGGGATTTTCCAAATTTCAAGTACAGTATCAATTAATCTTAATGCA 2034
DB 2781 AATATTAAGATTTTCAATCTTAATTAATTTTCTTAACAGATTAATTTGCTCCAAATCA 2840
```


QY	1203	TTTATGTTGGTATTTCTATTCGATGACCTACTTATATGCAATCTATTAACGTAAACCTT	1282
Db	2505	TTTACCAACGCATTATATATATGTTTCATTTACACACTTGATATATATATCCAAATATCTAG	2564
QY	1263	ATATGAGAAAGAAACAGGTTCAACCCACAACAAAAACAATAGACATTTGATCTTATTA	1322
Db	2565	TGTTTTGGAAATCAACAATGTAACTGATTAATTAATATCTCTGGTTTGGCAACAAATAT	2624
QY	1323	AGTTCAATTGTACTGATACACATACCTCCGTTTCCCTTATTCACACACCTTAT	1382
Db	2625	TTAATATTTTTTATTAATAATGTCATTAAGCTTAGAATTAATATATCTAAATATGATTTAATA	2684
QY	1383	AATTATCAAAATTTGAACCTTTATTTAAATGGCTCACTTAACAACAACACTCAATTTTCAGC	1442
Db	2685	TATTAGTAAATATGATTTTTTTTATTAACATAATGTACTGACTTTTGGAGAAAGAACTTAC	2744
QY	1443	AGGAGGGCT-----TTATCTAAATTATCAAAACAACATTTTTTCAATTTCTTAAGAA	1496
Db	2745	AGCAGGATCTGGGCAAAATACCTATATGATGTAAATATAAAATATTTTCGGTTACCAATCT	2804
QY	1497	AAAAGACTGCAATCTAGTATTTGATCCAGGTTGTTACGAAACTTTATATCTATATCTA	1556
Db	2805	TAAACGAAGAGAAATCAAGGAAACCCCTATTCCTTTTCCAAATATATGATTAATCTATAGCA	2864
QY	1557	TATTTTATCCCATTTTTCATATATTTACTTATTTCTATGTGATGATGATTAACAGTACAAAT	1616
Db	2865	TATTTTATC-----ATTATTAAGTCTTATATTCCTTGATATCCCTGCAACATAT	2907
QY	1617	ATTAGATACAGGTGTA--TTAGATGACACACACAGTATGTTGATAGATATATATGCAATA	1674
Db	2908	AAAACCTACAGGTATACGTTTGTCTTGGACACACACTCTAGTGTGATCTTAAATAATACATT	2967
QY	1675	TCAGATATAAATTAATCAATATGCCAGCATTCAAAGGTAACATCTTGATACAACTCT	1734
Db	2968	TATACACTTTTAACTACCCAAATTCACGCTGTAAAGGAAATTCACCTTGGGACGTCTCT	3027
QY	1735	AAGGTAATTGAGAGACCTGTCATACACAGAGAAACTTGTTTATTTTCAAAAGTCAAGG	1794
Db	3028	AAGGTTGTTCAAGGACCTGGTCATACAGAGAGGAGTTTAATTGATTTTCAAAAGATCA----	3083
QY	1795	CGTTTAGAATTTACATGTGAAACTCTTAATTTCTACACATCTTATTTTCATTTAGACTTGA	1854
Db	3084	--TTTCAAAATTAATCATGTCAACACTCAAAATTTTCAACATCTGTATTTTATTAAGAAATTCGT	3141
QY	1855	TATGCTCAAAATGCGTGAAGAAATACCTCTCTATATATCTCTTAACAATCCAGGAGTA	1914
Db	3142	TATGCTTCAAAATGGAAGCGCAAAATACCTGAGCTGTTAATAATCTTAGTATCCAGGG--	3199
QY	1915	ATAGAAATACCACTCAACGACTCAACAAACCTTTTCTGGTACAAATATATATATTTTA	1974
Db	3200	-TAGACAGACTGGGATATGCACTCAACCCCACTTTTCTGGTACAGATATATAGCAATTTTA	3258
QY	1975	CAATACGAGATTTTGGGATTTTCCAAATTTCCAAAGTACGTAACATTAACCTTTAATATGA	2034
Db	3259	AAATATATAAGATTTTTCAGTACTCTAGAAATTTCTAACGAGGGAATTTGCTCCAAATCA	3318
QY	2035	AACATACCATTTATATTTAATCGAGAGATGTAT--CAATTTCAATTTTAAATCATTTGAT	2091
Db	3319	AACATATCTCTTGTTTAAATCGTTTGGATGTATATATACAAACACACAGTACTTATTTGAT	3378
QY	2092	AAAAATGAAATTTATACCAATTAATCTCTATATGACCAACAAATATAGAAAAACAATA	2151
Db	3379	AAAAATGAAATTTCTGCAATTAATCTGTTCTATTAAGAGAGATATAGAGAAACAATAATTA	3438
QY	2152	GAAACTATCCAAACAAAAATTAATACATTTTTCACAAATCATACAAAA	2200
Db	3439	GAAACAGTACAAACAATATTAATTAACATTTTATGCAAAATCCTATTAATAAA	3487

RESULT 7
AAQ14669
ID AAQ14669 standard; DNA; 3543 BP.

XX	AAQ14669;
AC	
DT	27-AUG-2003 (revised)
DT	25-MAR-2003 (revised)
DT	04-FEB-1992 (first entry)
XX	
DE	Dipteran active toxin gene.
XX	
KM	Insecticide; B.t.; crystal; delta endotoxin; cryIIA; ss.
OS	Bacillus thuringiensis serovar morrisoni.
FH	Key
FT	CDS
FT	Location/Qualifiers
XX	1..3543
XX	/+tag= a
PN	EP457498-A.
XX	
PD	21-NOV-1991.
XX	
PF	09-MAY-1991; 91BP-00304180.
XX	
PR	15-MAY-1990; 90US-00524255.
PR	01-OCT-1990; 90US-00590903.
XX	
PA	(MYCO) MYCOGEN CORP.
XX	
P1	Sick AJ;
DR	WPI; 1991-341902/47.
XX	
DR	P-PSDB; AAR14373.
XX	
PT	Bacillus thuringiensis genes encoding diptera-active toxins - and
PT	transformed microbes used to control insects in various environments.
PS	Claim 1; Page 10; 20pp; English.
CC	The sequence was obt'd. from plasmid pMYC1625 which was isolated from a
CC	genomic library prep'd. from DNA from B.t. PS71M3 [from B.t. PS71M3-69
CC	(NRRL B-18515)]. It is related to the cryIVA family of genes, the 140 kD
CC	endotoxin gene and the type II gene from B.t. var. israelensis. The gene
CC	encodes a 110 kd protein. Microorganisms transformed with the DNA may be
CC	administered to dipteran insects or their environments, the expressed
CC	toxins acting as an insecticide. See also AAQ14670-Q14672. (Updated on 25
CC	-MAR-2003 to correct PA field.) (Updated on 27-AUG-2003 to correct OS
CC	field.)
XX	
SQ	Sequence 3543 BP; 1278 A; 567 C; 612 G; 1086 T; 0 U; 0 Other;
	Query Match 15.4%; Score 340.6; DB 2; Length 3543;
	Best Local Similarity 53.1%; Pred. No. 6.5e-50;
	Matches 1170; Conservative 0; Mismatches 934; Indels 99; Gaps 17,
OY	10 AATPACGATATTAAGCATATGAATTATTGATTCGCATACCTTCACTTATTTCCGAAC 69
Db	4 AATCTTATCAATAATTAATAATGAATATGAACAATTAAATGGTTTCAAAAAAATTAAAT 63
OY	70 AGAAAACGTAATGATTTGTAGATACCCTTACACAAATTAATCCAATCAACATTACAAAAC 129
Db	64 ATATCTAATATTAATTAACAAGTATCCAAATGAAATATGTCACAAAACAAATTTATCAAAGT 123
OY	130 ACAAAATTACAAAGTGCGCTCAATATNTGTCAAGGGAATACAAATATGGGATTAATTC 189
Db	124 ACAAAATTATAAGTTGGCTCAATATNTGTCAACAGAATACAGCATATGTGTGAGATTTT 183
OY	190 GAGACATTTGCTAGTGCTGATACAAATTGCTGCAGTTAGTGACAGGTACTATTGTCCGGT 249
Db	184 GAACATTTTATTGAT-----AGTGTGAATCACTAGTGCTATACATTGTATGTTGGG 234
OY	250 ACTCTGTTAGCCGGTATATAGTGGGCTACTTTATATCCGGAACCGATNAGAAATATAGST 309
Db	235 ACCGTACTGACTGT-----TTGGGTTTCAACAACCTTTAGA 273

QY 310 GCTATATATATCTTTGGTACCCCTATCACTGCTTTTGGCCGCGGAGACAAAGAC 369
 DB 274 CTGCTTATATAGTTTGGTACATTAATACAGAGTTCTTTTCCAGCCCAAGACCAATCT 333
 QY 370 AAAACAGTATGACACAAATTTATTAATATGGAGAAATTTTGTGATACCCGTTTACA 429
 DB 334 AACAC---ATGAGAGTACTTTATTAACAACAATAAATATTTATTAATAAATAATAGCA 390
 QY 430 -GAAAGCATTAACAGCTTAAGTTTCAACCTTAGAGAGTTTACCAATATTTACAAG 488
 DB 391 TCAACATATATATATGCTATATTAATTTTAAACAGTCTGTTATATGTTATCAGCACT 450
 QY 489 CTATATATCAGAGTTAGATGATGGAGAAATTTAAAGATACACAGCTCTGATTTACC 548
 DB 451 TATCATATATCCTTTAAACATGGGAGAAATATCCAAACCCCAAAATATCTGAGATGTA 510
 QY 549 ACCATCATCAGCATTAACAACAAGTGCCTTGACTCTTAATAATACGATTTGAGATGTTCA 608
 DB 511 AGGA----CACAAATCCAGCTAGTTTATTAACATTTTCAAAATGTCAATTCAGACCTTGT 566
 QY 609 CAATGATTTTATTCAGAGAAATACCTGTTTCCAACTGTAACCTTATTAACAGCTATTACT 668
 DB 567 AAACCTTGTCTCTCTAATCTAGTATGTC-----GATTAATAACATATCTATGAT 620
 QY 669 ACCATTTATGCGCAAGCTGCTAATTTTCAATTTTAAATTTATTAACAACAGTCTGATTT 728
 DB 621 ATCTAGTTATGCAACAAGCAGCAAACTTATCATCTGATCTGATTTAAATCAACGCTGCAAT 680
 QY 729 GCGTATGATGAGAAATGCAATATATCATCTTCACAAAATGGAACCTTAATGCTGAAATC 788
 DB 681 TGAACGATTTTAAATAAACAATGCAATTCGATTTATTTAGAGCTTGGCC---AACAGC 737
 QY 789 AGATGACTATTTAATCTTTTAAAGAAATATACCTAATATAGTAACTATTTGGCA 848
 DB 738 AATTATATATATCCATATGACTTAAGCTATTAAGATTAAGATTAACATTAATTTTGTATAC 797
 QY 849 TACCTATATGACAGAGCTTAAATAATCTTGAAGACCAACCAATATGAAATGAGATATTT 908
 DB 798 AACTTATATTAAGATTTAAATTTTAAACGACGCTGATATGATATCTTGAATGAGAA 857
 QY 909 TAATGACTATGAGATATATGACCATTACTGTATTTAGATACATCTCTCAATTTTCTT 968
 DB 858 TATAAATCTGAACATTA-----CAATGCTATGCAACAAATATGACTACTGCTATTT 911
 QY 969 ATATGATATTAATAAATATATAGATTTCAATAGAGAAATGAGAAATGAGATTAAGAA 1028
 DB 912 AGATCTTGTGCACTCTTCTTAATTTATGATGATAGTAAATATCCAAATAGGTGTCCAATC 971
 QY 1029 TGAATCTCAAGAGAAATTTATTAACAAGTGAATTAATTTGATGCTTCTCAACTTAG 1088
 DB 972 TGAATCTTACTGAGAAATTTAT--CAGGTACTTAATCTTGAAGAAAGCCCTATTAATAT 1029
 QY 1089 AGTTCAACCCATCTAGCTACAGTGAATATTAATTTAAACAGCTGCAAGTTTAAATTTAT 1148
 DB 1030 TATGACTTTCAATATCAAGAGATTTCACTTACAGTAAGAC--GCATTTATTTACTGGCC 1087
 QY 1149 TTGATTTTGAACAATTTATTTTATTAACAGAAATATCAAAATTTGCGGAATGTTTACT 1208
 DB 1088 TTGATTTCTTGA-----TTTATTAAGAAAGCGAAATCTCTTAATTAATTTTTCAC 1142
 QY 1209 TGGTATTTCTAATCGATGACACTACTTATAGCAATCTAATACTGAATCTTTATATATG 1268
 DB 1143 CAGCATATATATATGTTTCACTTACACACTGATATATATATCCAAAATCTAGCTTTT 1202
 QY 1269 AGAAAGAACAGGTTACCCACAACAAACAAATTAAGACATTTGATTTTAAAGTTTC 1328
 DB 1203 TGAATATCAATGTAATGATTAATTAATCTCTGTTTGGCAACAATAATTTATAT 1262
 QY 1329 AATTGTAATGATAGCAATACCTCTGTTTCCCTATTTCAACGACCTTATATATTA 1388
 DB 1263 TTTTATTAATATGCTATAGTCTTATGATTAATTAATTTCAATTAATTAATTAATG 1322

QY 1389 TCAATTAAGCTTTATTTAAATGCTCATCTTACAAACACTCAAAATTTACAGACAGG 1448
 DB 1323 TAAATGATTTTATTTATTAATATGATAGTACTTTTGGAGAAAGACTTACAGACAG 1382
 QY 1449 GTCT-----TTACTATTAATCAAAACAACTTTTTCATTTCTTCTAGAAAAAGA 1502
 DB 1383 ATCTGGCAAAATTAATTAATGATTAATTAATTAATTTTGGGTATCAATTTCTTAAAC 1442
 QY 1503 CTGCAATCTAGTTATTTGATTCAGAGTTGTTCAACCAACTTTAATTAATTAATTAATTT 1562
 DB 1443 AAGAGAAATCAAGAAATCCCTACCTTTTCCAAATATATATATATATGATATTTT 1502
 QY 1563 ATCCCATTTTTCATTTATTTACTTATTTCTATGATGATTAAGTATTAACGTACAAATTTGA 1622
 DB 1503 ATC-----ATTATTAAGTCTTATATCTTACCTGCAACATATTAATCT 1545
 QY 1623 TACAGTGTAT--TTAGATGACACACAGTGTGTTGATATATATATGCAATATCAAT 1680
 DB 1546 CAAGTATATAGTTTGTGCTGACACACTCTAGTGTGATCCTAATAAATACATTTATACA 1605
 QY 1681 AAAATTAATTAAGATGATCCAGCAATCAAGGTACAAATCTTGAATACAAAGCTTAAGGTA 1740
 DB 1606 CATTTACTACCAAAATTCAGCTGTATTAAGCAATTTCACTTGGGATGCTTTCTTAAGTT 1665
 QY 1741 ATGAAAGACCTGTGCTATACAGAGAAACTTGTGTTTATTAACAAGTCAAGGCGTTTA 1800
 DB 1666 GTTCAAGACCTGGTCAATACAGAGAGGATTTAATGATTTCAAAAGATCA-----TTTC 1719
 QY 1801 GAAATTAATGATGAAATCTCTAATTTCTACAAATTTTCAATTTTCAATTTGATGATGCT 1860
 DB 1720 AAAATTAATGATGCAACACTCAAAATTTTCAAAATGATTTTATTAAGAAATTCGTATGCT 1779
 QY 1861 ACAAATGCTGGAATAATCTCTCTAATATATCTCTTAACAAATTCAGAGATTAATGGA 1920
 DB 1780 TCAATATGAAAGCAAAATACAGAGCTTATTAATCTTATGATCCAGGG--TAGCA 1836
 QY 1921 ATACCACTCAACGACTCAACAAACCTTTTCTGTATCAAAATTAATTAATTTTCAATAC 1980
 DB 1837 GAACTGGGATATGGAACCTCAACCCCACTTTTCTGGTACAGATTAATGCAATTTAAATAT 1896
 QY 1961 GGAATTTTGGGATTTTCAATTTTCAAGTACAGTACATTAATCTTAAATGGAACATA 2040
 DB 1897 AAAGATTTTCAATCTTGAATTTTCTTAACGAGGAAATTTGCTCCAAATCAAAACATA 1956
 QY 2041 CCATTTATTTAATTCGAGATGAT--CAATCAATTTTATGATGATTAATTT 2097
 DB 1967 TCTCTGTTTAAATCGTTGGAATGATATTAACAACACACACAGTACTTATGATTAATTT 2016
 QY 2098 GAATTTATACCAATTAATCTCTATGACACCAAAATAGAGAAAAAATTAAGAACT 2157
 DB 2017 GAATTTCTGCCAATTAATCTGTTCTATTAAGAGATTAAGAGAAACAAATTTAGAAACA 2076
 QY 2158 ATCCAAACAAATTAATTAATTAATTTTCAAAATCAATCAAAAAA 2200
 DB 2077 GTACAAACAAATTAATTAATTAATTTATGCAATCTATTAATAA 2119

RESULT 8
 AA081178
 ID AA081178 standard; DNA; 3543 BP.
 AC AA081178;
 AC 25-MAR-2003 (revised)
 DT 12-AUG-1995 (first entry)
 DE B.c. toxin P871M3 gene.
 DE Delta-endotoxin; crystal protein; biological control agent; Calliphoridae;
 KW screw-worm; sheep blowfly; Lucilia; Phormia; Calliphora; insecticide;
 KM pesticide; B.c.; ss.
 XX
 OS *Bacillus thuringiensis*.

```
XX W09502694-A2.
XX
XX 26-JAN-1995.
XX
XX 13-JUL-1994; 94MO-US007902.
XX
XX 15-JUL-1993; 93US-00093199.
XX
XX (MYCO ) MYCOGEN CORP.
XX
XX HICKLE LA, Payne J;
XX
XX WPI, 1995-067338/09.
XX P-P8DB; AAR63078.
XX
XX Method for controlling Caliphoridae pests - specifically utilizing
XX Bacillus thuringiensis isolates or toxins.
XX
XX Disclosure; Page 36-38; 50pp; English.
XX
XX A library was constructed from Bacillus thuringiensis PS71M3 total
XX cellular DNA in lambda Gem-11. Plasmid pMYC1625, selected in Baculovirus
XX coli, contained a 8.0 kb insert expressing a beta-endotoxin gene. This
XX was sequenced (AA081178). A cured, acrycetaliferous B.c. host carrying
XX pMYC1625 produced a 130 kDa crystal protein (AAR63078) toxic to
XX caliphorids. (Updated on 25-MAR-2003 to correct PN field.)
XX
XX Sequence 3543 BP; 1278 A; 566 C; 612 G; 1087 T; 0 U; 0 Other;
XX
Query Match 15.4%; Score 340.6; DB 2; Length 3543;
Beet Local Similarity 53.1%; Pred. No. 6.5e-50;
Matches 1170; Conservative 0; Mismatches 934; Indels 99; Gaps 17;
XX
XX 10 AATAACGATTAATGCAATATGAAATTTGATTCGACCTACCTGATTTTCCGAC 69
XX 4 AATCTTATCAAAATGAAATGAAATGAAATGAAATGAAATGAAATGAAATGAAAT 63
XX
XX 70 AGAAACGATTAATGATTTGATTAACCTTACACAAATTAATCCAAATCAACATTA 129
XX 64 ATATCTAATTAATTAATCAAGATATCAATAGAAATAGTCCAAACAAATTAATCA 123
XX
XX 130 ACAATTAACAAAGATGCTGCTCAATATGCTCAAGGAATACAAATATGATTAATTC 189
XX 124 ACAATTAATTAAGATGCTGCTCAATATGCTCAAGGAATACAAATATGATTAATTC 183
XX
XX 190 GAGACATTTGCTAGTGTGATACAAATGCTGCAAGTATGCAAGGATTAATTCGGT 249
XX 184 GAACTTTTATGAT-----AGTGTGAACCTGAGTGTCTATCTATTTGATTTGG 234
XX
XX 250 ACTCTGTAGCCGATTAAGTGGGCTCACTTCTATATCCGAGATAGAAATATAGT 309
XX 235 ACCGTACTGACTGT-----TTCCGGTTCAACAACCTTATGGA 273
XX
XX 310 GCTAATTAATATCTTTTGTGTAACCTTAATCACTGCTTTTGGCCCGGGAACAAGAC 369
XX 274 CTTCCTTAATTAAGTGTGTAATTAATTAATCAAGTCTTTTTCAGCCCAAGACCAATCT 333
XX
XX 370 AAAACAGTATGCAACATTTATTAATTAATGGAAGAAATTTTGTGATACACGTTAACA 429
XX 334 AACAC--ATGAGATGACTTTTATACACAACTAAATATTTATTAATTAATTAATTA 390
XX
XX 430 -GAAAGCATTAAGAGCTTAAGATTAACAACTTTAGAAAGATTTAGACAAATATTA 488
XX 391 TCAACATTAATTAAGTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 450
XX
XX 489 CTATTAATTAAGATTAAGTATTAATTAATTAATTAATTAATTAATTAATTAATTA 548
XX 451 TATCATTAATTAAGTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 510
XX
XX 549 ACCATCATCAGCATTAACAAGAGCTGCTGACTCTTAAATTAAGATTTGAAATGTTCA 608
XX 511 AGGA-----CACAAATCAAGCTAGTTCATTCACATTTTCAAAATGTCAATTCAGAGCTTGT 566
```

```
QY 609 CAATGATTTTATTCGAGAAATACCTGGTTTCAACTTGAACTTATTAACGCTATTA 668
DB 567 AAACCTTGCTCTCTAATCTAGTATGTC-----GATTAATTAACATTAATTA 620
QY 669 ACCATTTATTCGCAAGCTGCTAATTTTCAATTAATTAATTAATTAATTAATTA 728
DB 621 ATCTAATTAATTCGCAAGCTGCTAATTTTCAATTAATTAATTAATTAATTAATTA 680
QY 729 GAGTGAATTAAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 788
DB 681 TGAAGCTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 737
QY 789 AGATGCTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 848
DB 738 AATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 797
QY 849 TACCTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 908
DB 798 AACTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 857
QY 909 TAATGCTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 968
DB 858 TATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 911
QY 969 ATATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1028
DB 912 AGATCTTGCTGCTCTCTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 971
QY 1029 TGAATCTGCAAGAAATTTTATTAATTAATTAATTAATTAATTAATTAATTAATTA 1088
DB 972 TGAATCTGCAAGAAATTTTAT--CAGTCTTAATTAATTAATTAATTAATTAATTA 1029
QY 1089 AGTTCACCAATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1148
DB 1030 TATGATCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1087
QY 1149 TTCAATTTTGAACAAATTAATTTTATTAATTAATTAATTAATTAATTAATTAATTA 1208
DB 1088 TTGATCTTTGAA-----TTTTATGAAGAAAGGCAATCTGCTAATTAATTTTTCAC 1142
QY 1209 TGTATTTCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1268
DB 1143 CAGCATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1202
QY 1269 AGAAAGCAAGCTTCAACCAACAACAAATTAAGATTAATTAATTAATTAATTA 1328
DB 1203 TGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1282
QY 1329 AATTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1388
DB 1263 TTTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1322
QY 1389 TCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1448
DB 1323 TAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1382
QY 1449 GTCT-----TTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1502
DB 1383 ATCTGGCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1442
QY 1503 CTGCAATCTAGTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1562
DB 1443 AAGAGAAATCAAGCAATCCCTACCTTTTCCAAATTAATTAATTAATTAATTAATTA 1502
QY 1563 ATCCATTTTCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1622
DB 1503 ATC-----ATTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1545
QY 1623 TACAGGTGA--TTAGATGAGACACAGATGATTAATTAATTAATTAATTAATTAATTA 1680
DB 1546 CAAGTGTATGCTTTGCTTGGACACACTGATGATTAATTAATTAATTAATTAATTAATTA 1605
```

Qy	1661	AAAATAATTACAAATGATCCGAGCAATCAAAGTAACAATCTTGATACAAATCTTAAGTA	1740
Dp	1606	CATTAACTACCCAAATTTCCAGCTGTAAAMCGAATTCATCTGGAGCTGCTTTAAGGTT	1665
Qy	1741	ATTGAGGACCTGGTCATACAGGAGGAACCTGGTTATTTACAAAGCCAAAGGCGTTTA	1800
Dp	1666	GTCACAGGACCTGGTCATACAGGAGGAGTTTAATGATTTCCAAAGATCA-----TTTC	1719
Qy	1801	GAATTTACATGTGAAACCTCCTAATTTCTACAAATCTTTTCATTAGACTTCGATATGCT	1860
Dp	1720	AAATTTACATGTTCACAACTCAAATTTTCAACAAATCGATTTTATAGAAATTCGTATGCT	1779
Qy	1861	ACAAATGGTCTCGAATAATCTCTTCTTAATATCTCTTACAAATCCAGAGTAATAGA	1920
Dp	1780	TCAATATGGAACCGAAATAACAGAGCTGTATTAATCTTATATCCGAGGG---TAGCA	1836
Qy	1921	ATACCACTCAACGATCTCAACAACACTTTTCTGTACAAATTATTAATTTTCAATAC	1980
Dp	1837	GAACCTGGGTATGGACATCCACCCCACTTTTCTGTACAGTTATACAAATTTTAAATAT	1896
Qy	1981	GGAGATTTTGGGTATTTTCCAAATTTCCAAAGTACAGTAACATTACCTTTAAATCGAAACATA	2040
Dp	1897	AAAGATTTTTCGTACTTATGAAATTTTCTTACAGAGTGAATTTGCTCCAAATCAAAACATA	1956
Qy	2041	CCATTTATATTTAATCGTCAGATGAT---CAAATTCATTTTATCATATGATTAATTT	2097
Dp	1957	TCCTCTGTGTTTAATCGTTGGAGTATATATCAAAACACACAGATCTTATTTGATTAATTT	2016
Qy	2098	GAATTTATACCAATTTACTCTCTATATGCACCAAATATAGAGAAAACAAAATTTAGAACT	2157
Dp	2017	GAATTTTCGCACATTTACTCGTTCTATATAGAGAGATATAGAGAAACAAAAATTTGAAACA	2076
Qy	2158	ATCCAAACAAATATTAATACATTTTTCACAAATCATACAAAAA	2200
Dp	2077	GTCACAGCAAAATATTAATATCAATTTATGCAAAATCTCTATAAAAA	2119

XX	RESULT 9
XX	AAQ14670
XX	ID AAQ14670 standard; DNA; 2061 BP.
XX	AAQ14670;
XX	
DT	27-AUG-2003 (revised)
DT	25-MAR-2003 (revised)
DT	04-FEB-1992 (first entry)
XX	
DE	Dipteran active toxin gene.
XX	
KW	Insecticide; B.t; crystal; delta endotoxin; cryIIVC, ss.
XX	
OS	Bacillus thuringiensis serovar morrisoni.
XX	
XX	Key
FT	Location/Qualifiers
FT	1..2028
FT	/*tag= a
XX	
FN	EP457498-A.
XX	
PD	21-NOV-1991.
XX	
PP	09-MAY-1991; 91EP-00304180.
XX	
FR	15-MAY-1990; 90US-00524255.
PR	01-OCT-1990; 90US-00590903.
XX	
PA	(MYCO) MYCOGEN CORP.
XX	
PI	Sick AJ;
XX	
DR	WPI; 1991-341902/47.
XX	P-PSDB; AAR14374.

PT *Bacillus thuringiensis* genes encoding dipteractive toxins - and
 PR transformed microbes used to control insects in various environments.
 XX
 PS Claim 1, Page 15, 20pp; English.

CC The sequence was obt'd. from plasmid pMYC1636 which was isolated from a
CC genomic library prep. from DNA from B.t. PS71M3 [from B.t. PS71M3-69
CC (NRRL B-16515)]. It is related to the cYIVC from B.t. var. israelensis.
CC The gene encodes a 77 kD protein. Microorganisms transformed with the DNA
CC may be administered to dipteran insects or their environments, the
CC expressed toxins acting as an insecticide. See also A014659-Q14672.
CC (updated on 25-MAR-2003 to correct PA field.) (updated on 27-AUG-2003 to
CC correct OS field.)

SQ Sequence 2061 BP; 779 A; 331 C; 319 G; 632 T; 0 U; 0 Other;

Query Match	8.5%	Score 188.2;	DB 2;	Length 2061;
Best Local Similarly	50.0%;	Pred. No. 1.9e-23;		
Matches 1107; Conservative	0;	Mismatches 923;	Indels 185;	Gaps 18;

OY	1	ATGAAATCAAAATACGATATATACGAATATGAAATATATGATTCGCATACCTCACTTAT	60
Db	1	ATGATTCATATCAAAATATAGAAATATATGAAATATTCAAATGCTC-----CATCAAT	54
OY	61	TTTCGGAACAGAAACAGTATATGATCTGTGATACCTCTTACCAAAATATCCAAATCAACCA	120
Db	55	GGTTTACCAAGTCTAATATACCTATTTCTGATATCCATTTAGCAAAATAGCCAAATCAACCA	114
OY	121	TTACAAAACACAAATTTACAAAGAGTGGCTCAATATGTGTCAAGGAAATACAAATATGCT	180
Db	115	CTGAAAAACAGAAATTTACAAAGATTTGGCTCATGTGTGTCAAGATATATCAACAATATGGC	174
OY	181	GATTAATTTGAGACATTTGCTATGTGCTGATACAAATTCGTCTGACGTTATGTGAGTACTAT	240
Db	175	AATAATCGGGGAAATTTGTGTATGTTCTGAACTAATGTGTGGAGTTATGTCAGGTAATTAAT	234
OY	241	GTATCCGCTACTCTGTATGACCGGTATAGGTGGGCTCACTTATATTCGGGACCGATATGGA	300
Db	235	GTATGAGAACTATTTGTATGAGCTTTGTCTGCCCT-----270	
OY	301	ATAATAGGTGCTATATATATATCTTTTGGTACCTTAATCACTGTCTTTTGGCCCGGGGA	366
Db	271	GTCTTAGCTGACGATATATATCTTTGGGACTTTGTGCCATCTTTTGGCAAGATCT	330
OY	361	GAAACAGACAAAACAGTATGACACAAATTTATTAATATGGGAAATTTTGTGTATACA	420
Db	331	GACCCCTG---CAAAATGTTTGGACGATTTGTAAACATGCGAGGAAAGCCCTATACAAGA	387
OY	421	CGTTTAAACAGAAAGCATATAAAACAGCTAAAGTTAAGTAAACCTTTGGAAGGATTTGACAAAT	480
Db	388	ATGATATATAAATCATATATATATGTACTTACTT-----CTATGATACCTATATAAAAAT	441
OY	481	TTACAAAGCTATATATACAGCATTTAGATATTTGGAGAAATTTAAAAAGACTACAGCTCCT	540
Db	442	CAACTGTATATATATATCAAGAAATTTTTCATTAATGGGAGCCAGCAGTACACACGCTAAT	501
OY	541	GGATTTACCAACATATCTACGACATTTACAAAGAGCTGCTTCACTCTTAAATATGATTTTGA	600
Db	502	GCTAAAGAGATATCAATGATCTCTTATCACTTAGAACCTATATATGATTAAGATTTTGA	566
OY	601	AATGTTCACAATGATTTTATTTGAGAAATATCTGTGTTCCAACTTGAAACTTATTAAGG	666
Db	561	TATGTTAAAA-----ATAATGCTATGTGCAAT	591
OY	661	CTATTTACTACTTATTTATGCGCAAGCTCTAATTTTCATTTAAATTTATTTACAAAGGT	720
Db	592	CCAACACTCCCTGCAATATGACAAATATGCTACTTGSCACTTGAAATTTATTAACAAATGCT	651
OY	721	GCTGAATTTGGCTGATGAAATGGAATGCAATATATCAATCTTCACAAAATTTGAACCTATGCT	780
Db	652	GCTACCTATTTACAAATATATGCTGCAAAATCAAGGAT-----AAATCAACAGTACTTTC	705
OY	781	GAACATGAGATGACTATTTATTAACCTTTTAAAAAGAAATATATCACTAATATATGATATCTAT	840

```

Db      706 AATTCACTTAATTAATCTATCAGGGCTATTATTAAGCGTAATAACAGTAATATACGACTAT
Qy      841 TGTGCAATATACCTATAGAACAGAGACTAATAAAATCTTAGAGAGAACCAATATGAAATGG
Db      766 TGTATACAAAGCTACATAGCAGAGACTAATCTAGATTAGAACCTAATATCTACAGCAACATGG
Qy      901 AGTATATTTAATGACTATGAGATATATGACATTTACTGTATATGATACATCTTCAA
Db      826 AATATGTATATTAATCTTACCGTTTGAATAATGACTTAATCTGTATAGATCTTATTTGCTAT
Qy      961 TTTTCTTATATGATATTAATAAGATATAGATTTCAATAGAGAGAAATGAAATGAAAGGC
Db      886 TTTCCAAATTAATGACCCGAAAAAATATCC-----AATAGGA 921
Qy      1021 ATTAAGAATGAACTACACAGAGAAATTTATACAACTGA-AATTAATTTTGTCTCTCC
Db      922 GTTAATATGAACTTACACAGAGAAATTTATAGAAATGTTAATTCAGATACATTTAGAAC
Qy      1080 TCAACTAGAGTCAACCCAACTAGCTAGACATGAGATATTAATTTACACGTGCAAGTTT
Db      982 ATTAACAGAACTAGAAAAATGATTAAGTAACTGAAATCTTACATTTATTACTTGGATTAACCA
Qy      1140 TAAATTAATTTCAATTTTATGAAACAATTTATTTTATACAGAAAAATACAAATTTGCGGAA
Db      1042 GGGCGTTTTCACAGAGAAATTTTCGAGACATCTTGATTCCTTATGATATTTTCTTTT
Qy      1200 TCGTTTATGTTGGTATTTCTTAATCGTATGACCTACTATATGACAACTATATACGAAAC
Db      1102 ACAGGTAAACAG-ATGGCTTTTACACATATGATATGACAACTAATCTGGGAGC 1160
Qy      1260 TTTATATGAGAAAGAACAGGTTCACCCACAAACAAATTAAGACATTTGAATCTTA
Db      1161 GGTTCATGACATATTTATTTCTCAAGACATCCAA-----GTAATTTCC 1205
Qy      1320 TAAAGTTCAATTTGAATGATGACATGACATCCTCTGTTCCCTATTAACACACTT
Db      1206 TTTTATATGAAACAAACCTATTTGATAGGTGAAATTTTCAGACATGAGATGACTCAGA
Qy      1380 TATAATTAATCAAAATTTGAATTTTATTAATGCTCATCTAACACACACTCAAAATATTC
Db      1266 TATAATATATGAATATATTTTTCGATAGC-----AGTGAAGATTTTCATATTC 1319
Qy      1440 AGCAGAGAGGCTTTATCTAATATCAAAACAACTTTTTCATTTCTAGAAAAA 1499
Db      1320 ATCCAAATTCACAAATAGAAATTAATTAATAAGAACTAATCTTATATGATTCACAAACA
Qy      1500 AGACTGCAATCTAGTATTTGATCCAGTTGTTCCACCAACTTAATACTATAGCATAT
Db      1380 AACATG-----GAAAAATTAAGAAATATGTCATAC 1409
Qy      1560 TTTATCCATTTTCAATTAATTAATTAATCTGATGATGATGATTAACAGTACAAATATTT
Db      1410 TCTATCGTATATTAATAAAGCTGATATTA-----ATATTTTCAGATGATAGGAAG 1460
Qy      1620 AGATACAGGTGTATAGATGACACACAGATGTTGTATGATATTAATGCAATATTCAGA
Db      1461 AAGAAGAGTTCATTTAGTTGACACATATAGTGTATTTCCAAATATACAAATAGATTT
Qy      1680 TAAATATTAATTAATGATCCAGCAATCAAAAGTAAATCTTGATATAAACTGTAAGT
Db      1521 AGATACATCAACCCAAATCCAGCTCTTAAGCTTGAAGGTGAAGTTGTGTAAGAAAT
Qy      1740 AATTGAAGAGCTGTGCTATACAGAGAGAACTTGTGTTATTTTACAAAGTCAAGGCGTTT
Db      1581 TGTGAAAGTCTGTGCTACACAGGTGAGACTTGTATTTCTTAAGATATGATAGATTT
Qy      1800 AGAAATTAATCATGTGAAACTCTTAATTTTACACAAATCTTATTTCAATGACTGATATGC
Db      1641 TAGAGTTAGATTTTAAAAAATGTTTCTGACAAATATACATAGACGTA---TTCCGTTATGC
Qy      1860 TACAATATGTCTGAAATATCTCTCCATATATATCTTACAAATCCAGAGATTAATAGG

```

```

Db      1698 TACTAATGCTCCAAAGCAACAGT--ATTCTTAACCGGAATAGATACTATTAAGTGGAG
Qy      1920 AATACCACTTCAACGACTCAACAACTTTTCTGTATCAAAATTAATTAATTAACATA
Db      1756 CTCCTAGTACCACTTCCCGCAAAAC-----CAATGCTACAGATTTTAACATA
Qy      1980 CGAGATTTTGGGTATTTCCAAATTTCCAGTACAGT-----AACATTTACTTTAA
Db      1806 TGCAATTTTGGATATTAATTAATTTCCAAAGACAGTTCCAAATTAACATTTGAAGAGA
Qy      2031 TCGAATATCACTTATTAATTTAATGCGTATGACCAAAATCATCTATATATATATGA
Db      1866 AGACACTTTATTAATGACCTTAATGCTTATGACCAAAATCATCTATATATATATGA
Qy      2091 TAAATTAATTAATTAATCAATTTCTCTATGACCAAAATAGAGAAAAACAAAAAT
Db      1926 CAAATATGAATTTATTTCAATCACTCAATCTGATTAATTAATACGAGAAAGAAATAT
Qy      2151 AGAACTATCCAAACAAAAATTAATATACATTTTTCACAAATCATCAAAAAACATT
Db      1986 AGAAAAACACAGAAAAATAGTGAATGATTAATTTGTTAATTAACAAAGTTCTT

```

RESULT 10

AA081180 ID AA081180 standard; DNA; 2061 BP.

XX AA081180;

XX AC AA081180; (revised)

XX DT 25-MAR-2003 (first entry)

XX DT 12-AUG-1995 (first entry)

XX XX B. l. toxin PS71M3 gene.

XX KM Delta-endotoxin; crystal protein; biological control agent; Calliphorid;

XX KM screw-worm; sheep blowfly; Lucilia; Phormia; Calliphora; insecticide;

XX KM pesticide; B. l.; ss.

XX OS Bacillus thuringiensis.

XX XX W09502694-A2.

XX PD 26-JAN-1995.

XX XX 13-JUL-1994; 94WO-US007902.

XX XX 15-JUL-1993; 93US-00093199.

XX PA (MYCO) MYCOGEN CORP.

XX PI Hickie LA, Payne J;

XX DR WPI; 1995-067338/09.

XX DR P-PSDB; AAR63079.

XX PT Method for controlling Calliphoridae pests - specifically utilizes

XX PT Bacillus thuringiensis isolates or toxins.

XX PS Disclosure; Page 42-43; 50pp; English.

XX CC A library was constructed from Bacillus thuringiensis PS71M3 total

XX CC cellular DNA in Lambda Gem-11. Plasmid pMTC1636, selected in Escherichia

XX CC coli, contained a 15 kb insert expressing a beta-endotoxin gene. This was

XX CC sequenced (AA081180). A cured, acrycetaliferous B. l. host carrying

XX CC pMTC1636 produced a 77 kDa crystal protein (AAR63079). (Updated on 25-MAR

XX CC -2003 to correct PW field.)

XX SQ Sequence 2061 BP; 779 A; 331 C; 319 G; 632 T; 0 U; 0 Other;

Query Match 8.5%; Score 188.2; DB 2; Length 2061;
 Best Local Similarity 50.0%; Pred. No. 1.9e-23;
 Matches 1107; Conservative 0; Mismatches 923; Indels 185; Gaps 18;

Db 1986 AGAAAAACACAGAAATAGTGAATTTATTGTTAATTAACAAAGTTCTT 2040

RESULT 11
AAN50525
ID AAN50525 standard; DNA; 3756 BP.
XX
AC AAN50525;
XX
DT 24-OCT-2003 (revised)
DT 25-MAR-2003 (revised)
DT 23-OCT-1991 (first entry)
XX
DE Bacillus thuringiensis var. israelensis endotoxin insert in plasmid
DS pS1367.
XX
KM Endotoxin; insecticide; ss.
XX
OS Bacillus thuringiensis serovar israelensis.
XX
PN EP153166-A.
XX
PD 28-AUG-1985.
XX
PF 15-FEB-1985; 85EP-00301017.
XX
PR 22-FEB-1984; 84US-00582506.
PR 22-JAN-1985; 85US-00693556.
XX
PA (SYTR) SYNTRO CORP.
XX
PI Walfield AM, Pollock TJ;
DR WPI; 1985-211724/35.
XX
PT Polypeptide active against Diptera insects - prepd. from DNA sequence
PT coding for Bti endotoxin using bacterial host.
XX
PS Disclosure; Page 15a-e; 27pp; English.
XX
CC The B. thuringiensis var. israelensis endotoxin insert in pS1367 is
CC expressed in a bacterial host. The protein produced has insecticidal
CC activity against dipteran insects. (Updated on 25-MAR-2003 to correct PA
CC field.) (Updated on 24-OCT-2003 to standardise OS field)
XX
SQ Sequence 3756 BP; 1379 A; 566 C; 659 G; 1150 T; 0 U; 2 Other;

Query Match 8.1%; Score 178.2; DB 1; Length 3756;
Best Local Similarity 50.0%; Pred. No. 11e-21;
Matches 1108; Conservative 0; Mismatches 923; Indels 186; Gaps 19;

QY 1 ATGAATCAAAATPACGATATATACGAATATGAATTTATGATTCGATACCTCACTTAT 60
DB 941 ATGATCCATATCAAAATATAGATATGAAATATTCATTCATC-----CATCCAT 994
QY 61 TTTCCGAACAGAAACGATATGATTTCTAGATACCTTACACAAATATATCCAAATCAACA 120
DB 995 GGTTTTACGAGTCTAATATATCTATCTAGATATCATATAGCAAAATATAGCCAAATCAACA 1054
QY 121 TTACAAAACAAATTTACAAAGAGTGCATATATGTCAGAGGAATATACCAATATAGT 180
DB 1055 CTGAATAACAGAAATTTACAAAGATGGCTCAATGTGTGCAAGATATCAACAATATATGC 1114
QY 181 GATTAATTTGAGACATTTGCTAGTGTCTGATACATATTCGTCAGTTAGTACGATATTT 240
DB 1115 AATATATGGGGGAATTTTGTCTAGTTCTGAACTATTTGAGGATAGTACGATATTTT 1174
QY 241 GTATCCGCTGATCTGTAGCCGGATATAGTGGGCTCACTTATATCCGACCGATAGCA 300
DB 1175 GTAGTAGAACAATATGTATGGAGCTTTGTGCCC-----CT 1210
QY 301 ATTAATAGGCGATATATATCTTTTGTACCCATCATCTGTCTTTTGCCCGCGGGA 360

Db 1211 GTCTTACCTGACGATATATATCTTTTGGACCTTTGTGCCGATCTTTTGGCAAGATCT 1270
QY 361 GAACAGACAAAACGATATGACACAAATTTATTAATTAATGGAGAAATTTTGTGATACA 420
DB 1271 GACCTG---CAATGTTTGGCAGAGATTTGTTAAACATCGAGAGAGCCATATCAAGAA 1327
QY 421 CCGTTAACAGAAAGCATTAATAACAGCTTAAAGTATCAAACTTTAGAGAGATTTAGCAATA 480
DB 1328 ATGATATAAAACATTAATTAATGTAATCTT-----CTATCGTACACCTATTAATAAT 1381
QY 481 TTACAAAGCTATTAATACAGCATTTAGATGATTTGGAGAAATTTAAAGACTACAGCTCCT 540
DB 1382 CACTGATTAATATATCAAGAAATTTTTCGATTAATGGAGCCAGACGATACAGCTAT 1441
QY 541 GGATTTACCAACATCATGACATTAACAACAGCTGCTTGACTTTAAATATGATTTGAG 600
DB 1442 GCTAAGACAGTA-CATGATCTCTTACTTACCTTGAACCTATATATGATTAAGATTTAGA 1500
QY 601 AATGTTACAAATGATTTTATTCGAGAAATACCTGTTCCAACTTGAACTTATTAACG 660
DB 1501 TATGTTAAAA-----ATATGCTAGCTATCGATA 1531
QY 661 CTATTACTACTATTTATGCGCAAGCTGCTAATTTTCATTTAAATTTATTAACAACAGT 720
DB 1532 CCAACACTCCCTGATATGCACAAAATAGCTACTTGGCACTTGAAATTTATTAACATGCT 1591
QY 721 GCTGAATTTGCTGATGATGGAATGCAATATACATCTTCACAAAATGAACTTAATGCT 780
DB 1592 GCTACCTATTAACAATATATGCTGCAAAATCAAGATTAATCCAACTATCTTCAAT-- 1649
QY 781 GGAACTACAGATGCTATTTATTAACCTTTAAAGAAATATCTTAATATAGTAACTAT 840
DB 1650 ----CATCTAATTAATCTATCAGGGCTATTTAAAGTAAATACAAAGATATATCTACTAT 1705
QY 841 TGTCAATATACCTATAGAACAGACTAATAAATCTTAGAGACGAACCAATATGAATG 900
DB 1706 GTATTAACAACGTAATGACAGACTAATGATTTGAACTAATATCTTAACGCAACTGG 1785
QY 901 AGTATATTTAATGACTATCGAAGATATATGACATTACTGTATTAATACATCTCTCA 960
DB 1766 AATATGATATATACTTACCGTTTGAATATGACTTACTGTGTTAGATCTTATGCTATT 1825
QY 961 TTTTCTTATATGATATTAATAAGATATAGATTTCAATAGAGAAATAGAAATGAAGGC 1020
DB 1826 TTTCCAAATATATGACCCAGAAATATATCC-----AATAGGA 1861
QY 1021 ATTAAGATGAACTACAGAGAAATTTATACACTGAATTAATTTGATCGTCTCT 1080
DB 1862 GTTAATCTGAACTTATCAGAGAAATTTATACGAATTTA----- 1902
QY 1903 -----TTCAGTACATTTAGAACCATTAACAACTAAGAAATGATTAATCTAGAAATCC 1957
QY 1141 AATATATTTTCATTTTATGAACTTTATTTTATATACAGAAATATCAAAATTTGGGAAT 1200
DB 1958 ACATTATTTCTTGGAATTAACCAAGGGCGTTTATACCAAGAAATTC--TCGACACTT 2014
QY 1201 CGTTAGTGTATTTCTAATCGTATGCACTCTTATATGCAATATCTATATCGAAACT 1260
DB 2015 CTGATCTTATATATATTTTCTTTTACAGTATACCGATATGAGCTTTTACATATCTAT 2074
QY 1261 TTATATGAGAGAAAGACAGCTTCAACCAACAACAAATTAAGACATTTGAATCTTAT 1320
DB 2075 GATGATGCAACATATATCTGGGAGCGGTTTATGGAATATTTATTTCTCAAGACATATCC 2134
QY 1321 AAGTTTCAATTTGATCTGATAGACATCACTCTGTTTCCCTATTTCAACACATTT 1380
DB 2135 AAGATATTTCTTTTATGAAACAAACCTATATAGGTGGAATTTG--CAGACATA 2192
QY 1381 ATTAATATCAAAATGAACTTATTTAATGAGCTCATCTAACAACACATCAATATTTCA 1440
DB 2193 GAGAGTACTCAGATATATATATGAAATGATATTTTTCGAATAGCAGTAAATTTT- 2251

1441 GCAGAGGCTCTTATCTAATTATCAAAAACAACTTTTTCATTTCTTAGAAAAA 1500
 2252 -----CGATTTTCATTCATCAACATAGAAAATATTAATAAGAACT----- 2296
 1501 GACTGCATCTAGATTATGATCCAGGTGTTCCAACTTTAATAATAGTCATATT 1560
 2297 -----GATTCCTATATGATTCCTAAAAACAACATGAAAAAATAGATATGCTACT 2350
 1561 TTATCCATTTTTCATTTATTTACTATTTCTATGATGATGATACAGTACCAATATTA 1620
 2351 CTATGCTATATATAAACTGATATATAT-----ATATTTTCAGTATGATAGAGAA 2401
 1621 GATACAGGTGATTTAGATGACACACAGTATGATAGATATATATCAATATCAGAT 1680
 2402 AGAAGGTGATTTAGTTAGTACACATCTAGTGTGATTTCCAAATATCAATAGTTTA 2461
 1681 AAAATATATTAATGATATCCAGCAATCAAGGTAACTATGATCAAACTCTAAGGTA 1740
 2462 GATTAACATACCCCAATCCACGCTTAAATAAAGTAAAGTTCTAATCAAAAAAT 2521
 1741 ATTGAAGACCTGTCATACAGAGAACTGTTTATTTCAAAAGTCAAGGCTTTA 1800
 2522 GTGAAGGCTGCTGTCACAAAGTAAACTTGTATTTCTAATAAGATGATGAAATTT 2581
 1801 GAAATTAATGATGAACTCTAATTTCTACACAACTTTATTTAGATGATGATGCT 1860
 2582 AGAGTTAGTTTATTAATAATGTTCTCAACATATCAAGTACGTA---TTGTTATGCT 2638
 1861 ACAAAATGCTGGAATACTCTTCTATATATCTCTTACAAATACAGAGTAAATGGA 1920
 2639 ACTAATGCTCCAAAGCAACAGTA--TTCTTAACGGAGATAGTATGTAAGTCTGGAAC 2696
 1921 ATACCACTCAACGACTCAACAACTTTTCTGAGCAAAATTAATTAATTTACATAC 1980
 2697 TCCCTAGTACCACTTCCGCAAAAC-----CCAAATGCTACAGATTTAATCAT 2746
 1981 GGAGATTTTGGGATTTTCCAACTTTCCAACTACAGT---ACATTTACCTTTAATGGA 2036
 2747 GCGATTTTGGATTTGATACATTTTCAGAACAGTTTCAATTAACATTTGAAGAGAA 2806
 2037 CATACATTTATATATTA-----ATCGTGAGATGATCAAAATCAATTTTAATCAT 2088
 2807 GACACTTTATTAATGACCTTATATGTTATGATACCAAAATCATTTATTAATATATAT 2866
 2089 GATTAATTTGAATTTTATACATTTCTCTCTATGACCAAAATAGAGAAAAACAATA 2148
 2867 GACAAATTTGAATTTATTCATCACTCAATCTGATTTAGATTAACAGAGAAACAAAT 2926
 2149 TTAGAACTATCCAAACAAATAATATACATTTTCAAAATCATTCAGAAAAACATT 2205
 2927 ATAGAAAAACACAGAAATAGTGAATGATTAATTTGTTAATTAACAAAGTTCTT 2983

RESULT 12
 AAD43974
 ID AAD43974 strand: DNA; 4896 BP.
 AAD43974;
 13-DEC-2002 (first entry)
 DE Bacillus thuringiensis ssp. finitimus cry28a1 gene.
 XX Delta-endotoxin; cry26a1; cry28a1; insect-resistant plant; toxin;
 XX transgenic host cell; insecticide; gene; ds.
 OS Bacillus thuringiensis.
 FH Key Location/Qualifiers
 FT CDS 1129..4458
 FT /tag= a
 FT /product= "cry28a1 gene"

XX US2002038005-A1.
 PN 28-MAR-2002.
 XX 08-JAN-2001; 2001US-00756526.
 PF 07-JAN-2000; 2000US-0175158P.
 XX (WOJC/) MOJICHOVSKA J A.
 PA (LEWIT/) LEWITTIN E I.
 PA (ZALU/) ZALUNIN I A.
 PA (REV1/) REVINA L P.
 PA (CHES/) CHESTUKHINA G G.
 PI Mojicchowka JA, Lewitin EI, Zalunin IA, Revina LP;
 PI Chestukhina GG;
 DR WPI; 2002-403936/43.
 DR P-PSDB; AAE26353.
 PT Novel isolated delta-endotoxin nucleic acid molecules, cry26a1 and
 PT cry28a1 isolated from Bacillus thuringiensis finitimus, that encodes
 PT toxin active against insects, useful for controlling insects.
 XX
 PS Claim 1; Page 33-37; 42pp; English.
 CC The invention relates to isolated delta-endotoxin nucleic acid molecules,
 CC cry26a1 and cry28a1 isolated from Bacillus thuringiensis finitimus,
 CC that encode a toxin that is active against insects. The invention is
 CC useful for producing an insect-resistant plant, by introducing the
 CC nucleic acid molecule into the plant, where the nucleic acid is
 CC expressible in the plant in an effective amount to control an insect. The
 CC invention is useful for producing a toxin that is active against insects
 CC by obtaining the transgenic host cell and expressing the nucleic acid
 CC molecule in the host cell, which results in the toxin that is active
 CC against insects. The toxin is useful for controlling an insect by
 CC delivering to the insect an effective amount of toxin. The invention is
 CC useful for controlling insects. The toxin is useful for inhibiting the
 CC ability of insect pest to survive, grow or reproduce, for limiting insect
 CC -related damage or loss in crop plants, and to prophylactically treat
 CC insect susceptible areas or plants to confer protection or resistance
 CC against harmful insects. The present sequence is Bacillus thuringiensis
 CC ssp. finitimus cry28a1 gene
 XX
 SQ Sequence 4896 BP; 1727 A; 757 C; 887 G; 1525 T; 0 U; 0 Other;
 Query Match 7.4%; Score 163.8; DB 6; Length 4896;
 Best Local Similarity 55.9%; Pred. No. 3.4e-19;
 Matches 405; Conservative 0; Mismatches 272; Indels 48; Gaps 3;
 QY 306 AGTGTCTATATATATCTTTTGTGATCACTGCTCTTTTGGCCGCGGAGAA 365
 DB 1293 AGTGTGTACTTATATATATGGAACCTTGCCGCTCTTTTGGCTGATCCAGAGA 1352
 QY 366 AGA---CAAAACAGATGACACATTTATTAATGGAAGAAATTTTGTGATACAC 422
 DB 1353 AGATCAAAAAAATTTGTGTCACATTTATGAAACAGAGAAACCTTTTAATCAAAAC 1412
 QY 423 GTTACAGAAAGCATTAACAGCTTAAGTTACAACTTTAGAGGATTTAGCAATAT 482
 DB 1413 AATTCTACAGCTGTAAGAAATATGATGATGATGATGATTTTAAAGATGAT 1472
 QY 483 ACAAACTATATATACAGATTTAGATTTGAGAAATTAATAAGATCAACAGCTCTG 542
 DB 1473 AACGTACTATGAAAGAGATTTAATGATTTGAAAGAAATCCAAATGCAAAATACGCG 1532
 QY 543 ATTACCAATCATCAGCATTAACAACAGCTGCTTGACTTTAATAATACGATTTGAGA 602
 DB 1533 ATTGTATACAGA-----GATTTGAAAA 1556
 QY 603 TGTTCACATGATTTTATTCGAGAAATACCTGTTTCCAACCTTGAACTTAATAACGCT 662

Db 1557 CGCTCATTTCAATTTTGTAGCAATATGCCACACTCCAACTTCCACGATAGACACTT 1616
Qy 663 ATTAACCTATTTATGCGCAAGCTGCTAATTTTCAATTTAAATTTATTAACAAGATGC 722
Db 1617 ATTATTAAGTTGCTATACAGAAAGCTGCAATTTTACATTTGATTTATTAATCAAGGTG 1676
Qy 723 TGAATTTGCTGATGAATGAATGCGAGATATACATCTTCCAAATTTGAACCTAATGCTGG 782
Db 1677 ACAATTCGCGATCATATGAAATGCGAGATCAACCAATTCACCAAT------GTTGAA 1727
Qy 783 AACATCATGATGCTATTAATTAATTTAAAGAAAATATACCTAAATATAGTAATTTG 842
Db 1728 GTCATCAGATCTATTTATTAAGACGCTATTTGTATATTTTGAAGATTTATTTATTTG 1787
Qy 843 TGCATAATCTATATAGACAGGACTTAAATTTTATAGACGACCAATATGAATGAG 902
Db 1788 CACCAAGACATACATTAAGATTTGATCCTTTAAAGATCAAGAAAATCATGCGA 1847
Qy 903 TATATTTAATGACTATGCAAGATATATGACCATTACTGTATTAATGATTCATCTCAAT 962
Db 1848 TGCTTATTAACATATTCGTCGAAATGACCTTAATTTGTATGATCTTGTGCGAACTTT 1907
Qy 963 TTTCTTATATGATATTAATAAAGATATAGAGATTTCAATAGAGAAATGAAGTAAAGCAT 1022
Db 1908 TCTTTTATATATATATACGTCGTTTCCAGAGAGTAACTAGAAATTAACAGAGAGGT 1967
Qy 1023 TAAGA 1027
Db 1968 TTATA 1972

RESULT 13
ADFI301
ID ADFI301 standard; DNA; 4896 BP.
XX ADFI301;
AC
XX 12-FEB-2004 (first entry)
DT
XX
DB Bacillus thuringiensis serovar finitimus pf2 DNA clone.
DE
XX Cry26Aa1; cry28Aa1; delta-endotoxin; insect pest control;
KW transgenic plant; insect resistance; insecticide; gene; ds.
XX
OS Bacillus thuringiensis serovar finitimus.
XX
OS
XX Key Location/Qualifiers
FH 1129.4458
FT CDS
FT /*tag= a
FT /product= "Cry28Aa1 delta-endotoxin protein"

XX US2003150018-A1.
XX
XX PD 07-AUG-2003.
XX PF 15-JAN-2003; 2003US-00345020.
XX PR 07-JAN-2000; 2000US-0175158P.
XX PR 08-JAN-2001; 2001US-00756526.
XX
XX PA (WOJC/) WOJCIECHOWSKA J A.
XX PA (LEWIT/) LEWITIN E I.
XX PA (ZALUN/) ZALUNIN I A.
XX PA (REVIT/) REVINA L P.
XX PA (CHES/) CHESTUKHINA G G.
XX
XX PI Wojciechowska JA, Lewitin EI, Zalunin IA, Revina LP,
PI Chestukhina GG;
XX
XX MPI; 2003-897623/82.
XX P-PSDB; ADFI3102.
XX
XX New isolated nucleic acid molecule encoding a toxin that is active

PT against insects useful for controlling insect pests or for conferring
PT insect resistance in plants.
XX
XX PS Claim 7; SEQ ID NO 3; 42pp; English.
XX
XX The present invention relates to the isolation of novel cry26Aa1 and
CC cry28Aa1 delta-endotoxin genes from Bacillus thuringiensis ssp.
CC finitimus. The sequences for the delta-endotoxin polypeptides are also
CC disclosed. The invention provides methods for producing the toxins and
CC compositions containing the toxins. The methods and sequences of the
CC invention are useful for controlling insect pests in transgenic plants to
CC confer insect resistance. The present sequence represents a DNA clone
CC that contains the coding sequence for cry28Aa1 delta-endotoxin.
XX
SQ Sequence 4896 BP; 1727 A; 757 C; 887 G; 1525 T; 0 U; 0 Other;
Query Match 7.4%; Score 163.8; DB 10; Length 4896;
Best Local Similarity 55.9%; Pred. No. 3.4e-19;
Matches 405; Conservative 0; Mismatches 272; Indels 48; Gaps 3;
Qy 306 AGGTGCTATATATATATCTTTTGTACCTAATCACTGTCTTTTGCCCGGAGAGACA 365
Db 1293 AGGTGTGCTATATATATCAATTTGGAACCTTGCGTCCGTTCTTTGGCTGATCCAGAGA 1352
Qy 366 AGA---CAAAACGATATGACACAAATTTATTAATGGAGAAATTTTGTGATACACC 422
Db 1353 AGATCCAAAATAATTTGTGCACAAATTTATGAACACGAGAGAACCTTTTAATCAAAAC 1412
Qy 423 GTTAACAGAAAGCATTAACAGCTAAAGTTACAACTTTTGAAGATTTTGAACAATATT 482
Db 1413 AATTTCTACAGCTGTAAAGAAATGCAATTTAGCTATCTTAATGTGTTTAAAGATGATT 1472
Qy 483 ACAAGCTATATATACAGCATTTAGATGATGAGAGAAATTAATAAGACTACAGCTCTGG 542
Db 1473 AAGCTACTATGAAGAGACATTTATGATGGAAGAGAAATCAAGTCAATATCTGCAG 1532
Qy 543 ATTAACCAATCATCAGCATTTACAAAGAGCTGCTTACTTTAAATAGATTGAGAA 602
Db 1533 ATTTGTATCACAGA-----GATTTGAAA 1556
Qy 603 TGTTCACAAATGATTTTATTTGAGAAATACCGTTTCCAACTTGAACCTTATTAAGAGCT 662
Db 1557 CGCTCATTTCAATTTTGTAGCAATATGCCACACTCCAACTTCCACGATAGACACTT 1616
Qy 663 ATTAACCTATTTATGCGCAAGCTGCTAATTTTCAATTTAAATTTATTAACAAGATGC 722
Db 1617 ATTATTAAGTTGCTATACAGAAAGCTGCAATTTTACATTTGATTTATTAATCAAGGTG 1676
Qy 723 TGAATTTGCTGATGAATGAATGCGAGATATACATCTTCCAAATTTGAACCTAATGCTGG 782
Db 1677 ACAATTCGCGATCATATGAAATGCGAGATCAACCAATTCACCAAT------GTTGAA 1727
Qy 783 AACATCATGATGCTATTAATTAATTTAAAGAAAATATACCTAAATATAGTAATTTG 842
Db 1728 GTCATCAGATCTATTTATTAAGACGCTATTTGTATATTTTGAAGATTTATTTATTTG 1787
Qy 843 TGCATAATCTATATAGACAGGACTTAAATTTTATAGACGACCAATATGAATGAG 902
Db 1788 CACCAAGACATACATTAAGATTTGATCCTTTAAAGATCAAGAAAATCATGCGA 1847
Qy 903 TATATTTAATGACTATGCAAGATATATGACCATTACTGTATTAATGATTCATCTCAAT 962
Db 1848 TGCTTATTAACATATTCGTCGAAATGACCTTAATTTGTATGATCTTGTGCGAACTTT 1907
Qy 963 TTTCTTATATGATATTAATAAAGATATAGAGATTTCAATAGAGAAATGAAGTAAAGCAT 1022
Db 1908 TCTTTTATATATATATACGTCGTTTCCAGAGAGATGAATTAACAGAGAGGT 1967
Qy 1023 TAAGA 1027
Db 1968 TTATA 1972

```
RESULT 14
ADF31306
ID ADF31306 standard; DNA; 4896 BP.
XX
AC ADF31306;
XX
DT 12-FEB-2004 (first entry)
XX
DE Bacillus thuringiensis serovar finitimus pfz DNA clone.
XX
KM Cry26Aa1; delta-endotoxin; insect pest control;
XX transgenic plant; insect resistance; insecticide; gene; ds.
XX
OS Bacillus thuringiensis serovar finitimus.
XX
FH Key Location/Qualifiers
FT 1129..4458
FT CDS /*tag= a
FT /product= "Cry28Aa1 delta-endotoxin protein"
XX
XX US2003154510-A1.
XX
XX 14-AUG-2003.
XX
XX 15-JAN-2003; 2003US-00342821.
XX
XX 07-JAN-2000; 2000US-0175158P.
XX
XX 08-JAN-2001; 2001US-00756526.
XX
XX (WOJC/) WOJCIECHOWSKA J A.
XX PA (LEMI/) LEWITIN E I.
XX PA (ZALU/) ZALUNIN I A.
XX PA (REV1/) REVINA L P.
XX PA (CHES/) CHESTUKHINA G G.
XX
XX PI Wojciechowska JA, Lewitin EI, Zalunin IA, Revina LP;
XX PI Chestukhina GG;
XX
XX WPI: 2003-897757/82.
XX
XX P-PSDB; ADF31307.
XX
XX New delta-endotoxin nucleic acid molecules, cry26Aa1 and cry28Aa1, useful
XX PT for controlling insect pests and for conferring insect resistance.
XX
XX PS Claim 7; SEQ ID NO 3; 42pp; English.
XX
XX The present invention relates to the isolation of novel cry26Aa1 and
XX CC cry28Aa1 delta-endotoxin genes from Bacillus thuringiensis ssp.
XX CC finitimus. The sequences for the delta-endotoxin polypeptides are also
XX CC disclosed. The invention provides methods for producing the toxins and
XX CC compositions containing the toxins. The methods and sequences of the
XX CC invention are useful for controlling insect pests in transgenic plants to
XX CC confer insect resistance. The present sequence represents a DNA clone
XX CC that contains the coding sequence for cry28Aa1 delta-endotoxin.
XX
XX SQ Sequence 4896 BP; 1727 A; 757 C; 887 G; 1525 T; 0 U; 0 Other;
XX
XX Query Match 7.4%; Score 163.8; DB 10; Length 4896;
XX Best Local Similarity 55.9%; Pred. No. 3.4e-19;
XX Matches 405; Conservative 0; Mismatches 272; Indels 48; Gaps 3;
XX
XX 306 AGGTGCTATTAATATCTTTTGGTACCTTAATCATCTGCTTTTGGCCCGGAGAAACA 365
XX |||||
XX DB 1293 AGGTGTGTACTTATATCATTTGGAACCTTGCTCCCTTCCTTGGCTGATCCAGAGA 1352
XX |||||
XX OY 366 AGA---CAAAACAGATGAGACAAATTTATTAAATGGAGAAATTTTGTGTATACAC 422
XX |||||
XX DB 1353 AGATCCAAAAAATTTGTGTCATTTATGAAACAGGAGAGACCTTTTAAATCAAAAC 1412
XX |||||
XX OY 423 GTTAAACAGAAAGCATTAACAGCTTAAAGTTACAAACCTTTAGAGAGATTAGCAAAATTT 482
XX |||||
XX DB 1413 AATTTCTACAGCTGTAAAGAAATGACATTAAGCTCATCTAAATGTTTAAAGATGATTT 1472
XX |||||
```

```
OY 483 ACAAGCTATATACAGCATTTAGATGATGGAGAAAAATTAAAAAGACTACAGCTCTGG 542
DB 1473 AAGCTACTATGAAGAGATTTATATGATGGAGAGAAATCCAGTGCAAAATATGCGCAG 1532
OY 543 ATTACACCATCATCAGCATTACAAACAAGCTGCTTGACTTTAAATAGCATTTGAGAA 602
DB 1533 ATTGTATCACAGA-----GATTTGAAAA 1556
OY 603 TGTTCACAAATGATTTTATGGAGAAATACCTGTTTCCAACTTGAACTTAATAAGCT 662
DB 1557 CGTCAATTTCAATTTTGTAAAGCAATATGCCAAACTCCACCTCCACGATATGACACATT 1616
OY 663 ATTACTACCTATTTATGGCAAGCTGCTAATTTTCAATTAATTTATTTATACAAAGTGC 722
DB 1617 ATTATTTAGTTGCTATACAGAAAGCTGCAAAATTTACATTTGATTTATATGATCAAGGT 1676
OY 723 TGAATTGGCTGATGAATGGAATGCAATATATCATCTTCACAAATTGAACCTAATGCTGG 782
DB 1677 ACAATTCGCGATCAATGGAATGCAATCAACCATTCACCAAT-----GTTGAA 1727
OY 783 AACATCAGATGACTATTTATTAACCTTTTAAAAAGAAATTTACTTAATATAGTAATATG 842
DB 1728 GTCAATCAGTACTTATTTATGACGACTATTTGTTATATTTGAAAAATATTTATTTATG 1787
OY 843 TGCAAATACCTATATGAACAGACTAAATAATCTTAGAGACGAACCAATATGAATGAGAG 902
DB 1788 CACCAAGACATACATATTAAGATTTGAATCAGCTTTAAAGAAATCAAGAAATATCACTGGA 1847
OY 903 TATATTTAATGACTATGGAAGATATATGACCATTAAGTATTAATATCAATCTCTCAATT 962
DB 1848 TGCCTTATTAACATATGCTGAGAAATGACCTTAATGTATTTGATCTTGTGCAACTTT 1907
OY 963 TTTCTTATATGATATTAATAAGATATATAGATTTCAATAGAGAGAAATAGAAAGTAAAGCAT 1022
DB 1908 TCCCTTTTATGATATATGATGCTGTTTTCCAAAGAGATGAACCTAATTAACAAGAGAGCT 1967
OY 1023 TAAGA 1027
DB 1968 TTATTA 1972
XX
XX RESULT 15
XX ID ABK51132
XX ABK51132 standard; cDNA; 3690 BP.
XX
XX AC ABK51132;
XX
XX DT 30-JUN-2002 (first entry)
XX
XX DE cDNA encoding Bacillus thuringiensis insecticidal protein.
XX
XX KM Insecticide; transgenic; Coleoptera larvae; ss; gene.
XX
XX OS Bacillus thuringiensis.
XX
XX FH Key Location/Qualifiers
XX FT 187..3690
XX FT CDS /*tag= a
XX FT /product= "Insecticide protein"
XX
XX PN JP2002045186-A.
XX
XX PD 12-FEB-2002.
XX
XX PF 03-AUG-2000; 2000JP-00236140.
XX
XX PR 03-AUG-2000; 2000JP-00236140.
XX
XX PA (SDSB-) SDS BIOTEC CORP.
XX
XX WPI: 2002-356468/39.
XX
XX P-PSDB; AAU80281.
XX
```

PT A protein having insecticidal activity, a DNA encoding said protein, and
an agent and a method for preventing harmful organisms.

XX Claim 4; Page 17-18; 19p; Japanese.

CC This invention relates to a crystalline protein comprising a fully
CC defined sequence and the nucleotide sequence encoding this protein. The
CC protein of the invention is an agent for preventing harmful organisms
CC comprising *Bacillus thuringiensis* serovar galleriae SDB502, its mutant or
CC a microbe transformed by a DNA encoding the protein. This microbe can be
CC used to produce a protein containing the protein, or containing a protein
CC having insecticidal activity produced by the SDB502, its mutant or a
CC transformed microbe, a microbe which is transformed by using the above
CC DNA and produces the above protein having insecticidal activity, a plant
CC or a seed transformed by using the above DNA, and *Bacillus thuringiensis*
CC serovar galleriae SDB502 producing a protein comprising and producing a
CC protein showing insecticidal activity. The protein of the invention may
CC have insecticidal activity. The agent is used for preventing Coleoptera
CC larvae. This sequence represents the cDNA encoding the *Bacillus*
CC *thuringiensis* insecticide protein of the invention

XX Sequence 3690 BP; 1289 A; 599 C; 762 G; 1040 T; 0 U; 0 Other;

Query Match 4.5%; Score 98.4; DB 6; Length 3690;

Best Local Similarity 43.4%; Pred. No. 7.7e-08; Matches 602; Conservative 0; Mismatches 751; Indels 33; Gaps 2;

```
QY 612 TGATTTTATTCGAAATACCTGGTTTCACACTTGAATATTAATAAAGCTATTACTACC 671
DB 717 TGATTTTGTCTAAATCCATCTTTTGCATATCTGACAGAGATACCATTTATATC 776
QY 672 TATTATGCGCAAGCTGCTAATTTTCATTAAATTTATTAACAACAAGTCTGAATGGC 731
DB 777 AGTATATGCAACACAGAGATTTTACATTTCTATTTATACGAATGCTTCATTTTGG 836
QY 732 TGATGANTGAGATGCAATATACATCTTCACAATTTGAACCTAATGCTGAACATCGA 791
DB 837 AGCAGAGTGG-----GGATTTCACCCAGAGAAATTTTC 869
QY 792 TGACTATTATTAACCTTTTAAAGAAATATATCTAAATATAGTAATTTTGCATAATAC 851
DB 870 CACATTTTATATGTCAGTGACACGTCACGCCCATATCTCGATTTATTTGTAAAGTG 929
QY 852 CTATAGAACAGAGACTAAATAATCTTAGAGCAACCAATATGAAATGAGATATTTTAA 911
DB 930 GTATTAACACTGCTTAGATTAATTAATAAGTATACATGCTCAAGTTGGCTGAAGTATCA 989
QY 912 TGACTATGCAATATATGACCATTTACTGTATTTAGATATACATCTCTCAATTTTCTTATA 971
DB 990 CCAATTCGGAAGAGAAATGACATTACTGTATTTAGATTTAGTACGTTATTTCCAAACTA 1049
QY 972 TGATATATAAAGATATAGATTCATATAGAGAGATAGAAATTAAGGATTAAGAAATGA 1031
DB 1050 TGACACACGATACGATTCACATCGAAACACAGGCCCACTTACAGGAGAGTATACGA 1109
QY 1032 ACTCACAGAGAAATTTATACACTGAATATATTTTGTGATGCTTCTCTCAACTTGAAGT 1091
DB 1110 TCCAATAGTATTTTAAACAGAGAAACAAGTGTGATTTTGTAGCGGTGTGTCACTTAACAG 1169
QY 1092 TCAACCCAACTTACTAGTACGATGAAATATTAATTAACAAGTGAAGTTTAAATTTTTC 1151
DB 1170 TGATATTTCTTTTTCAGAAAGTGAAGGCTGTAAATTCGTTTCACACACTTATTTGATAT 1229
QY 1152 ATTTTATGAAACAATTTTATTTTATACAGAAATATCAAAATTCGGGAATCGTTAGTTGG 1211
DB 1230 ACTCAGTGAATTAATTAATTTTATACACAGAGCGGGCTTCCCTTGAATTAATACGAATA 1289
QY 1212 TATTTTCAATGATGACCTACT-----TATAGCAATACTATTAACCTGAACCTTTATA 1265
DB 1290 CCTTGAATATTTGGTGAAGACATTTCTATTAATATTAATAATAGATGCTCATCAGCATT 1349
QY 1266 TGGAGAAAGAACAGTTTACCCACAAACAAATTAAGACATTTTGAATCTTATTAAGT 1325
```

```
DB 1350 AGAAGTAATTTACGGTACGATTTACTTTAACAATAATCAAGTATTTATGCAATTA 1409
QY 1326 TTCATATGTAAGTATGATGACATCACTCTGTTTCCCTATTTCACACACTTATAT 1385
DB 1410 GGAATATCTTTCAGGTTGATGATTTAGGGGCGGATTTATGTAATTTACTACGACAGGTATA 1469
QY 1386 TAATCAATTTGAACCTTATTTTAAATGCTCATCTTAACAACACTCAAAATTTTCAGCAG 1445
DB 1470 TGAAGTTCGATGACGTAATGTTTACACTCTTGACAAAGAAATACAGATCAGATTCAGTTGG 1529
QY 1446 AGGCTCTTATCTAATATATCAAAACACACACTTTTTCATTTCTTGAATAAAGACTG 1505
DB 1530 AGGTTTACGATCTCAAAACACATACATATGCAAGTATGTACAAAATTTGCAATAC 1589
QY 1506 CAATCAATGATTTGATTCAGAGTGTTCACCAAACTTTAATCACTATATGATATTTATC 1565
DB 1590 GATTGATGAATCCCTCCAGAGATGAGCCACTTAGTAGAGGGATATGCCATGATTTATC 1649
QY 1566 CCAATTTTCAATTTTATTTACTTTCTATGTGATTTGATTCAGCTACAAATTTAGTAC 1625
DB 1650 TCAATATCCTCTTATTTCTTTTCTAAGAAATGCTAGTATGCTGTAGATATGCAATCT 1709
QY 1626 AGGATATTTAGATGACACACAGTATGTTGATATATATATGCAATATCAATAAAT 1685
DB 1710 CCTGTATTTGCTTGGACACATGCGATGCGGATGTTTCAAAATACAGTTTATTCAGATTA 1769
QY 1686 AATTACAAATATCCACCAATCAAAAGTAAACAATCTTGATACAAACTTAAAGTAAATGA 1745
DB 1770 AATTACTCAGATATCAGGTTGTAAAGGACATATTTAGTTTCAAGTACTACTGTTATTTAA 1829
QY 1746 AGAAGCTGTGATACAGAGAGAAACCTTGTTTATTTACAAAGTCAAGGGGCTTTAGAAAT 1805
DB 1830 AGGCTCTGGAATTTACAGAGAGCAATATCTTAAAGAAACAAAGTATGCTCGTTAGCTTA 1889
QY 1806 TACATGGAACCTCTAATTTCTACAACTTTATTTCACTAGACTTGATATGCTACAA 1865
DB 1890 TACTAGTGTCTGTAAATATCACATTTATACAAAGATATCGTCAAGAAATAGCTTATGC 1949
QY 1866 TGGTGTGGAATACTTCTCTATATATCTCTTACAAATACAGAGATTAATAGGAATACC 1925
DB 1950 TTCTACTACTAATCTTACAGCTTTTGTATTAACAATTTTCGAATCTCGCATTTACTATATA 2009
QY 1926 ACCTCAACGACTCAACAACCTTTTCTGTGTACAAATTAATTAATTTTACAAATCGGAGA 1985
DB 2010 TGTTAATAAATCCATGATTAAGGGATGATTTAAACATTTAATATACATTTGACTTACGAAC 2069
QY 1986 TTTTGG 1991
DB 2070 TATTTGG 2075
```

Search completed: December 11, 2005, 14:34:14
Job time : 1341 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 11, 2005, 13:36:36 / Search time 1134 Seconds
(without alignments)
11272.709 Million cell updates/sec

Title: US-10-783-417-1
Perfect score: 2208
Sequence: 1 atgaatcaataacgataa.....atcatcaataaacacttga 2208

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapect 1.0

Searched: 5883141 seqs, 28421725653 residues

Total number of hits satisfying chosen parameters: 11766282

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenBembl:*
1: gb_da:*
2: gb_in:*
3: gb_env:*
4: gb_com:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pr:*
9: gb_ro:*
10: gb_sts:*
11: gb_sy:*
12: gb_un:*
13: gb_vl:*
14: gb_hcg:*
15: gb_pl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2208	100.0	2208	6	CQ868305 Sequence
2	1903.8	86.2	2235	6	CQ868307 Sequence
3	1794.4	81.3	2085	6	CQ868309 Sequence
4	346.6	15.7	4451	6	I08083 Sequence 1
5	346.6	15.7	4451	6	I09103 Sequence 1
6	345.4	15.6	3543	1	BTTXOD1 Y00423 Bacillus th
7	345	15.6	4253	1	BACISRH4 D0078744 Bacillus th
8	345	15.6	4934	6	E01676 DNA sequenc
9	345	15.6	127823	1	BTBPTOXIS AL731825 Bacillus
10	222.6	10.1	6009	1	BTBPTOXIS AL731825 Bacillus
11	189.6	8.6	127923	1	BTBPTOXIS AL731825 Bacillus
12	188	8.5	3753	1	BACMSQB M12662 B. thuringie
13	184.2	8.3	4334	1	D0078744 Bacillus
14	178.2	8.1	3756	6	E00614 DNA encodin
15	171.6	7.8	1993	6	I03578 Sequence 1
16	166.8	7.6	4959	1	AF285775 Bacillus
17	163.8	7.4	4896	1	AF132928 Bacillus
18	157	7.1	3752	1	AB125059 Bacillus

19	143.8	6.5	3642	1	AB193814
20	133.6	6.1	3644	1	BTB251978
21	98.4	4.5	3435	1	AB089299 Bacillus
22	98.4	4.5	3504	6	BD133574 Protein h
23	98.4	4.5	3690	6	BD133575 Protein h
24	98	4.4	1738	6	I03580 Sequence 3
25	97	4.4	104992	14	AC005504 Plasmodu
26	97	4.4	169546	14	AC004157 Plasmodu
27	97	4.4	250421	2	AE014849 Plasmodu
28	97	4.4	349980	6	AX344555 Sequence
29	94.6	4.3	15548	6	AX347057 Sequence
30	94.6	4.3	133877	14	AC120883 Homo sapi
31	92.4	4.2	86826	2	PFMAL3P5 AL034556 Plasmodu
32	91.8	4.2	250029	2	AE014839 Plasmodu
33	91.8	4.2	313050	2	PPA929352 Plasmodu
34	90.8	4.1	3668	1	D88381 Bacillus th
35	90.4	4.1	8056	6	AX599046 Sequence
36	86.2	3.9	254050	2	PPA929358 Plasmodu
37	86	3.9	4660	6	CS083843 Sequence
38	86	3.9	90550	8	AL592166 Human DNA
39	86	3.9	176898	8	AC117569 Homo sapi
40	85.8	3.9	3444	6	AR050192 Sequence
41	85.8	3.9	3444	6	AR050194 Sequence
42	85.8	3.9	3444	6	AR060320 Sequence
43	85.8	3.9	3444	6	AR060322 Sequence
44	85.8	3.9	3444	6	I19463 Sequence 22
45	85.8	3.9	3444	6	I19465 Sequence 26

ALIGNMENTS

RESULT 1
LOCUS CQ868305 2208 bp DNA linear PAT 13-SEP-2004
DEFINITION Sequence 6 from Patent WO2004074462.
ACCESSION CQ868305
VERSION CQ868305.1 GI:5198351
KEYWORDS
SOURCE Bacillus thuringiensis
ORGANISM Bacillus thuringiensis
Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus; Bacillus cereus group.
REFERENCE 1
AUTHORS Carozzi,N., Hargis,T., Koziel,M.G., Duck,N.B. and Carr,B.
TITLE Delta-endotoxin genes and methods for their use
JOURNAL Patent: WO 2004074462-A 6 02-SEP-2004;
Athenix Corporation (US)
FEATURES
Source
1..2208
/organism="Bacillus thuringiensis"
/mol_type="unassigned DNA"
/db_xref="taxon:1428"
1..2208
/note="unnamed protein product"
/codon_start=1
/transl_table=11
/protein_id="CAH33948.1"
/db_xref="GI:5198352"
/translation="MNQNDNNEYEIIIDSHSPYFPNNNSDRPYTNNDPQLONT
NYKEMLMNCQNTGDNFEFTASADITAAVSAGTIVSGTLAGIGLTSISGPGLI
GALISFGLITVFWPAGBODKVTWGTGKIGELFVDPDLFESIKQLQLEGRQI
LOSNTLDDRRKIKRLQAPGLPPSALDQQAALTKIRFENVHNDPFIIEIGFOLEY
KTLILPIYAQANAHNLTLOCGALADBNADIDIPSOIEPPNAGSDVYKIKENIKP
YSNVCANTYRTGLKRLRDEPMKKSIRFNDYRRTITVLDITISQFSLDIKRYDSIG
GIEVAGIKNELTREIYVTEINFDRLPOLRVGPVNLATWENVTLRASFLEDEPIFY
TEENTNFGRLVGINRDAPTYNTITETLYGERTGSPTTKTIRFESYKVAIVDROS
PVSPFIOPHFIINOIEIYLNSSNNITLYKSAGSLSNYQNTTIRFQPKQCNVYIDP
GSPNPNVSHIILSHFSLFTYSYVIGLOILDVGIVGWTSSVDRAVADIKITMI
PAIKGNLIDNMSKVIKRGHGVGLVYLOSGRGREITCENSNQSVIRLRVANGA
GNTLPNLSLTIPVIGIIPPORLANTPSTRTNNQIGDPCGFGPSPSTYTLPLNRIFP
IFNRADVNSLILIDKLEIFPITSSMRQNRKQLETTIOTKINFPFNHTXTL"

ORIGIN

Query Match 100.0%; Score 2208; DB 6; Length 2208;
Beet Local Similarity 100.0%; Pred. No. 4,9e-250;
Matches 2208; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGAATCAAAATACGATATATACGAATATGAAATTTATGATTCGATACCTCACTTAT 60
DB 1 ATGAATCAAAATACGATATATACGAATATGAAATTTATGATTCGATACCTCACTTAT 60

QY 61 TTTCCGAAACAGAAACAGTATGATTTCTAGATACCTTTACACAAATATCCAAATCAACCA 120
DB 61 TTTCCGAAACAGAAACAGTATGATTTCTAGATACCTTTACACAAATATCCAAATCAACCA 120

QY 121 TTACAAAACAAATTTACAAAAGTGGCTCAATATGTCTCAAGGAAATACCAATATGGT 180
DB 121 TTACAAAACAAATTTACAAAAGTGGCTCAATATGTCTCAAGGAAATACCAATATGGT 180

QY 181 GATAATTTGAGACATTTGCTAGTCTGATACAAATTTGCTGACGTTAGTGCAGGTACTAT 240
DB 181 GATAATTTGAGACATTTGCTAGTCTGATACAAATTTGCTGACGTTAGTGCAGGTACTAT 240

QY 241 GTATCCGGTACTCTGTATAGCCGGTATAGTGGGCTCACTTATATCCGACCGATAGGA 300
DB 241 GTATCCGGTACTCTGTATAGCCGGTATAGTGGGCTCACTTATATCCGACCGATAGGA 300

QY 301 ATATAGGTGCTATATATATCTTTTGGTACCTTAATCACTGTCTTTTGGCCGGGGA 360
DB 301 ATATAGGTGCTATATATATCTTTTGGTACCTTAATCACTGTCTTTTGGCCGGGGA 360

QY 361 GAACAGACAAACAGTATGAGACAAATTTTAAATGGAGAAATTTTGGTATGATCA 420
DB 361 GAACAGACAAACAGTATGAGACAAATTTTAAATGGAGAAATTTTGGTATGATCA 420

QY 421 CCGTTAACAGAAAGCATATAACAGCTTAAAGTTACAACTTTAGAGATTTAGCAATA 480
DB 421 CCGTTAACAGAAAGCATATAACAGCTTAAAGTTAGCAACTTTAGAGATTTAGCAATA 480

QY 481 TTACAAAGCTATATATACAGCATTTAGATTTGGAGAAATTTAAAGACTACAGCTCT 540
DB 481 TTACAAAGCTATATATACAGCATTTAGATTTGGAGAAATTTAAAGACTACAGCTCT 540

QY 541 GGATTAACACCATCATCATGACATTTACACAGAGCTGCTTGAATTTTAAATTCGATTTGAG 600
DB 541 GGATTAACACCATCATCATGACATTTACACAGAGCTGCTTGAATTTTAAATTCGATTTGAG 600

QY 601 AATGTTCAAAATGATTTTATTCGAGAAATACCTGGTTTCCAATTTGAACCTTATPAAACG 660
DB 601 AATGTTCAAAATGATTTTATTCGAGAAATACCTGGTTTCCAATTTGAACCTTATPAAACG 660

QY 721 GCTGAAATTTGGGATGAGAAATGAGAAATGATATACATCTTCACAAATTTGAACCTTAATGCT 780
DB 721 GCTGAAATTTGGGATGAGAAATGAGAAATGATATACATCTTCACAAATTTGAACCTTAATGCT 780

QY 781 GGAACATCAGATGATCTATTTATTAATCTTTTAAAGAAATATATCTAAATATATAGTAACAT 840
DB 781 GGAACATCAGATGATCTATTTATTAATCTTTTAAAGAAATATATCTAAATATATAGTAACAT 840

QY 841 TGTGCAAAATACCTATAGAAACAGGACTTAAATCTTTAGAGAGAAACCAATATGAAATGG 900
DB 841 TGTGCAAAATACCTATAGAAACAGGACTTAAATCTTTAGAGAGAAACCAATATGAAATGG 900

QY 901 AGTATATTTAATGATCTATCGAAGATATATGACATTTACCTGATTAAGATCAATCTCTAA 960
DB 901 AGTATATTTAATGATCTATCGAAGATATATGACATTTACCTGATTAAGATCAATCTCTAA 960

QY 961 TTTTCTTTATATGAT 1020
DB 961 TTTTCTTTATATGAT 1020

QY 1021 ATTPAAGATGATCTCACAGAGAAATTTTATACATCTGAAATTTAATTTTGATTCGCTTCT 1080
DB 1021 ATTPAAGATGATCTCACAGAGAAATTTTATACATCTGAAATTTAATTTTGATTCGCTTCT 1080

QY 1081 CAACCTTAGAGTCAACCCCAATCTAGTACGATGGAATATATATTTTAAACAGTGCAGTTT 1140
DB 1081 CAACCTTAGAGTCAACCCCAATCTAGTACGATGGAATATATATTTTAAACAGTGCAGTTT 1140

QY 1141 AAATTTATTTTCAATTTTAAACAAATTTTATTTTATACAGAAATTTACAAATTTTCGGAAAT 1200
DB 1141 AAATTTATTTTCAATTTTAAACAAATTTTATTTTATACAGAAATTTACAAATTTTCGGAAAT 1200

QY 1201 CGTTAGTGGTATTTCTAATTCGATGACCTTACTTATATGCAATTAATTAATCTGAAACT 1260
DB 1201 CGTTAGTGGTATTTCTAATTCGATGACCTTACTTATATGCAATTAATTAATCTGAAACT 1260

QY 1261 TTATATGAGAAAGAAACAGGTTCCACCAACAAACAAATTAAGACCATTTGAATCTTAT 1320
DB 1261 TTATATGAGAAAGAAACAGGTTCCACCAACAAACAAATTAAGACCATTTGAATCTTAT 1320

QY 1321 AAAGTTTCAATTTGATCTGATAGCAATCACTCTGTTCCCTATTTCAACCACTTT 1380
DB 1321 AAAGTTTCAATTTGATCTGATAGCAATCACTCTGTTCCCTATTTCAACCACTTT 1380

QY 1381 ATATTTAATCAAAATGGAATTTTATTAATGAGTCTATCAACACACACTCAATATTTCA 1440
DB 1381 ATATTTAATCAAAATGGAATTTTATTAATGAGTCTATCAACACACACTCAATATTTCA 1440

QY 1441 GCAGAGAGGCTTTTATCTAATTTATCAAAACCAACTTTTTCATTTCTAGAAAAAA 1500
DB 1441 GCAGAGAGGCTTTTATCTAATTTATCAAAACCAACTTTTTCATTTCTAGAAAAAA 1500

QY 1501 GACTGCAATCTAGTATATGATTCAGGTTTTCACGAACTTTTATATATATGATCATAT 1560
DB 1501 GACTGCAATCTAGTATATGATTCAGGTTTTCACGAACTTTTATATATATGATCATAT 1560

QY 1561 TTATCCATTTTTCATTAATTTATCTTATTTCTATGATGATGATGATGATGATGATGAT 1620
DB 1561 TTATCCATTTTTCATTAATTTATCTTATTTCTATGATGATGATGATGATGATGATGAT 1620

QY 1621 GATACAGGTGATATAGATGGAACACAGTATGATGATGATGATGATGATGATGATGAT 1680
DB 1621 GATACAGGTGATATAGATGGAACACAGTATGATGATGATGATGATGATGATGATGAT 1680

QY 1681 AAAATATTTAATCAATGATCCAGCAATCAAAAGTAACTTTGATACAACTCTAAGGTA 1740
DB 1681 AAAATATTTAATCAATGATCCAGCAATCAAAAGTAACTTTGATACAACTCTAAGGTA 1740

QY 1741 ATGGAAGGCTGGTCAATACAGAGGAACTGGTTTATTTTCAAAAGTCAAAAGGCGTTTA 1800
DB 1741 ATGGAAGGCTGGTCAATACAGAGGAACTGGTTTATTTTCAAAAGTCAAAAGGCGTTTA 1800

QY 1801 GAAATTAATGATGAACTCTTAATTTTACAAATCTTATTTTCAATTAAGTCTGATAGCT 1860
DB 1801 GAAATTAATGATGAACTCTTAATTTTACAAATCTTATTTTCAATTAAGTCTGATAGCT 1860

QY 1861 ACAAAATGCTGGAAATATCTCTCTAATATATCTCTTCAAAATACAGAGGATATAGGA 1920
DB 1861 ACAAAATGCTGGAAATATCTCTCTAATATATCTCTTCAAAATACAGAGGATATAGGA 1920

QY 1921 ATACACACTCAAGACCTCAACAACTTTTCTGGTACAAATTTATTAATTTTCAATATAC 1980
DB 1921 ATACACACTCAAGACCTCAACAACTTTTCTGGTACAAATTTATTAATTTTCAATATAC 1980

QY 1981 GGAATTTTGGGTATTTTCAATTTTCAAGTACAGTATCAATTAATCTTTAAATTCGAAACATA 2040
DB 1981 GGAATTTTGGGTATTTTCAATTTTCAAGTACAGTATCAATTAATCTTTAAATTCGAAACATA 2040

QY 2041 CCATTTATATTTATTCGATGAGATGATCAATTTCAATTTTATCATTTGATATTAATTTGAA 2100
DB 2041 CCATTTATATTTATTCGATGAGATGATCAATTTCAATTTTATCATTTGATATTAATTTGAA 2100

QY 2101 TTTATACCAATTTACTTCTCTATGACCAAAATATGAGAAAAACAAATTTAGAAACTATC 2160

|||||
Db 2101 TTATATACCAATTACTCTCTATGACCAAAATAGAGAAAAAATAATAGAACTATC 2160
Qy 2161 CAACAAAAATAATATCATTTTTCACAAATCATATACAAAAACCTTGA 2208
Db 2161 CAACAAAAATAATATCATTTTTCACAAATCATATACAAAAACCTTGA 2208
RESULT 2
CQ868307 2235 bp DNA linear PAT 13-SEP-2004
LOCUS CQ868307
DEFINITION Sequence 8 from Patent WO2004074462.
ACCESSION CQ868307
VERSION CQ868307.1 GI:5198353
KEYWORDS
SOURCE Bacillus thuringiensis
ORGANISM Bacillus thuringiensis
Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus; Bacillus cereus group.
REFERENCE 1
AUTHORS Carozzi, N., Hargiss, T., Koziel, M.G., Duck, N.B. and Carr, B.
TITLE Delta-endotoxin genes and methods for their use
JOURNAL Patent: WO 2004074462-A 8 02-SEP-2004;
Athenix Corporation (US)
FEATURES
source 1..2235
/organism="Bacillus thuringiensis"
/mol_type="unassigned DNA"
/db_xref="taxon:1428"
1..2235
/note="unassigned protein product"
/codon_start=1
/protein_id="CAH33949.1"
/db_xref="GI:5198354"
/translation="MNQNNNEVEYIIDSKNLSYSPSNRIIDHSRYPYTNNPNOLNTN
YKEMLMOCQYQYDNEFTFASADTIAVSAAGTIVSTLAGIGLTSISGPIGIG
AIIISFGLTIVFWPAGBODKVTWTFPIKMGIIIVDTPLTSSIKOLKOTLEGPOIL
OSYNTALDWRKLRKLOAPGLPPRSALQOALITLKRPEVTHNDRIEIRPGOLETVK
TLUPRTYQAAANFHLNLOQCAKELDENRADIHPQIERNAGTSBDYTKLKENIPKY
SNYCANTRYREGLNKRNENPMKWSIFNDYRRYMTITVDITLAQSFYDIKRYKOSIGR
IGGIKTELRLREIYTEINFEDRLTYLEIOPNLAMEYNIIRSGLRISFLDELIVYTKN
ETYGRLVGIANRNSRYATGTEIIEYERGPPTTKLIPESYKSVIWDROVTP
SPFNITFIQINDELYLNNSPSNKLTYAGSGLSNDKKTDPQFPVKDCRIPINPC
LPSNYSVSHIISOPSLFNYSYKIGLALNIIYTGALGAMTHSSNPNNAISDKITIMPA
IKGNSLDNRSKRTIBSGHTGAGLVYLGQGRLETCRIPNSQSYTITLPRATNGAGN
TLPNISLTIPGVIGIPQRLNNTFSGTINNLQYDGFYFQPSVTLPILNRNIPFI
NRADVSNILIIIDKIEFIPIITSSVRONREKOLETITQIKINTFPNTHTKNTLNTBATN
YDID"

ORIGIN
Query Match 86.2%; Score 1903.8; DB 6; Length 2235;
Best Local Similarity 92.2%; Pred. No. 86-249;
Matches 2030; Conservative 0; Mismatches 162; Indels 9; Gaps 2;
11 ATAAAGATAATACGAATATGAAATTAATTGATTCGATACCTACCTTATTTTGGACA 70
Db 8 AAAATAATATATATATATATGATATATGATTTCAAGAAATTTATCTTATCTCTTACA 67
Qy 71 GAAACAGTAATGATTTCTAGATACCTTACACAAATATCCAATCAACCATTAACAAA 130
Db 68 GAAATATGATCATTTCTAGATACCTTACACAAATATCCAATCAACCATTAACAAA 127
Qy 131 CAATATCAAGAGTGGCTCAATATGTCGAGGGAATACCAATATGTCGTAATTTGG 190
Db 128 CAATATCAAGAGTGGCTCAATATGTCGAGGGAATACCAATATGTCGTAATTTGG 187
Qy 191 AGACATTTGCTAGTGTGATACATTTGCTGAGTAGTGAGTGTCTATGATCCGGTA 250
Db 188 AGACATTTGCTAGTGTGATACATTTGCTGAGTAGTGAGTGTCTATGATCCGGTA 247
Qy 251 CTCTGTAGCCGGTATAGTGGCTCACTTCTATATCCGACCGATAGGAATATAGGTG 310
|||||

Db 248 CTCTGTAGCCGGTATAGTGGCTCACTTCTATATCCGACCGATAGGAATATAGGTG 307
Qy 311 CTATATATATATCTTTGTGACCCCTAATCACTGTCTTTGGCCGGGAGAACAGACA 370
Db 308 CTATATATATATCTTTGTGACCCCTAATCACTGTCTTTGGCCGGGAGAACAGACA 367
Qy 371 AAACAGATGAGACAAATTTTAAATAATGAGAAATTTTGTGATACACGTTTAAACG 430
Db 368 AAACAGATGAGACAAATTTTAAATAATGAGAAATTTTGTGATACACGTTTAAACG 427
Qy 431 AAAGCATTAACAGCTAAAGTTTACAACTTTAGAAAGATTAGACAAATATTTACAAAGCT 490
Db 428 AAAGCATTAACAGCTAAAGTTTACAACTTTAGAAAGATTAGACAAATATTTACAAAGCT 487
Qy 491 ATATATGAGCATTTGATGATATTTGAGAAATTAATAAGCTCAAGCTCTGTGATTAACAC 550
Db 488 ATATATGAGCATTTGATGATATTTGAGAAATTAATAAGCTCAAGCTCTGTGATTAACAC 547
Qy 551 CATCATGAGCATTTACAAACAGCTGCTGACTCTTAAATAATGAGATTGAGAAATTTTACA 610
Db 548 CATCATGAGCATTTACAAACAGCTGCTGACTCTTAAATAATGAGATTGAGAAATTTTACA 607
Qy 611 ATGATTTTATTCGAGAAATACCTGTTTCAACTTGAACCTTATTAACAGCTATTAAC 670
Db 608 ATGATTTTATTCGAGAAATACCTGTTTCAACTTGAACCTTATTAACAGCTATTAAC 667
Qy 671 CTATTTATGCCCAAGCTGCTAATTTTCACTTAAATTTTAAATTAACAAAGCTGTGAATGG 730
Db 668 CTATTTATGCCCAAGCTGCTAATTTTCACTTAAATTTTAAATTAACAAAGCTGTGAATGG 727
Qy 721 CTGATGATGAGAAATGACAGATATACATCTTCAAAATTTGAACCTTATCTGGAATCAG 790
Db 728 CTGATGATGAGAAATGACAGATATACATCTTCAAAATTTGAACCTTATCTGGAATCAG 787
Qy 791 ATGATATATTAATCTTTTAAAGAAATATATCTTAAATATATGATATATTTGTGCAATA 850
Db 788 ATGATATATTAATCTTTTAAAGAAATATATCTTAAATATATGATATATTTGTGCAATA 847
Qy 851 CCTATATGAAAGCTAAATAAATCTTGAAGGAAACAAATATGAAATGAGATATATTA 910
Db 848 CCTATATGAAAGCTAAATAAATCTTGAAGGAAACAAATATGAAATGAGATATATTA 907
Qy 911 ATGATATGAAAGCTAAATAAATCTTGAAGGAAACAAATATGAAATGAGATATATTA 970
Db 908 ATGATATGAAAGCTAAATAAATCTTGAAGGAAACAAATATGAAATGAGATATATTA 967
Qy 971 ATGATATGAAAGCTAAATAAATCTTGAAGGAAACAAATATGAAATGAGATATATTA 1030
Db 968 ATGATATGAAAGCTAAATAAATCTTGAAGGAAACAAATATGAAATGAGATATATTA 1021
Qy 1031 AACTCACAAGGAAATTTATCAACTGAAATTAATTTGATCGTCTTCAACTTAAG 1090
Db 1022 AACTCACAAGGAAATTTATCAACTGAAATTAATTTGATCGTCTTCAACTTAAG 1081
Qy 1091 TTCAACCAATCTAGCTAGATGAGAAATTAATTTTAAACAGTGAAGTTTAAATATTTT 1150
Db 1082 TTCAACCAATCTAGCTAGATGAGAAATTAATTTTAAACAGTGAAGTTTAAATATTTT 1141
Qy 1151 CATTTTGAACAAATTTATTTTATACAGAAATACAAATTTTGGGAATCGTTTATGTTG 1210
Db 1142 CATTTTGAACAAATTTATTTTATACAGAAATACAAATTTTGGGAATCGTTTATGTTG 1201
Qy 1211 GTATTTTGAATCGTATGACCTTATATGCAATATCTATATCTTATATGAG 1270
Db 1202 GTATTTGCAATCGTATATGATCTTATATGTCAGACGAGAACTGAAATTTATATATGAG 1261
Qy 1271 AAAGAAAGGTTCCACCCCAACAAATAATGAGCAATTTGAATCTTATATGAGTTTCA 1330
Db 1262 AAAGAAAGGTTCCACCCCAACAAATAATGAGCAATTTGAATCTTATATGAGTTTCA 1321
Qy 1331 TTGTATCTGATATGACATCACTCTCTGTTCCCTATTC--AACCACTTTATATTA 1387
Db 1322 TTGTATCTGATATGACATCACTCTCTGTTCCCTATTC--AACCACTTTATATTA 1381
|||||

QY 1388 ATCAATTTGAATTTTAAATGGCTCATCTAAACAACACTCAAAATATTACAGAGAG 1447
DB 1382 ATCAAAATTTGAATTTTAAATTAATTCACCTAGTAATTAATTAACATATTACAGCTGGG 1441
QY 1448 GGTCTTATCTAAATTAATTAACAACTTTTTCATTTTTCCTAGGAAAAAAGACGTCA 1507
DB 1442 GGAATTTATCTAAATTAATTAACAACTGATTTTCATTTTCTGTAAAAAAGAGCTGA 1501
QY 1508 ATCTAGTTATTTGATCCAGTGTGTTCAACCAACTTAAATTAATTAATTAATTTTATCCC 1567
DB 1502 AACCAATTTATTAATTCAAATTTGTTTACCAAGCTAATTAATTAATTAATTTATCCC 1561
QY 1568 ATTTTTCATTATTTTAACTTATTTCTATGATGATTTGATTAAGCTTCAAAATTTAGATACAG 1627
DB 1562 AGTTTCTTTTATTTTAAATTAATTCCTAATAAATTTGATTTAGCCCTAATTAATTAATTAATCAG 1621
QY 1628 GTGTATTAAGATGACACACAGTAGTGTGATAGATTAATTAATTAATTAATTAATTAATTA 1687
DB 1622 GTGCTTATTAAGATGACACACAGTAGTGTGATTAATTAATTAATTAATTAATTAATTAATTA 1681
QY 1688 TTAAATGATCCAGCATCAAAAGGTAAACAATTTGATTAACAAGCTTAAGGTAAATTTGAG 1747
DB 1682 TTAAATGATCCAGCATCAAAAGGTAAACAATTTGATTAACAAGCTTAAGGTAAATTTGAG 1741
QY 1748 GACCTGTGATTAAGATGACACACAGTAGTGTGATTAATTAATTAATTAATTAATTAATTA 1807
DB 1742 GACCTGTGATTAAGATGACACACAGTAGTGTGATTAATTAATTAATTAATTAATTAATTA 1801
QY 1808 CATGTGAAATCTCTAATTTCTACACAACTTATTTCAATTAAGCTTCAATGCTCAATG 1867
DB 1802 CATGTGAAATCTCTAATTTCTACACAACTTATTTCAATTAAGCTTCAATGCTCAATG 1861
QY 1868 GTGCTGAAATTAATCTCTAATTAATTTCTTAACAATCAAGAGTAAATTAAGATTAACAC 1927
DB 1862 GTGCTGAAATTAATCTCTAATTAATTTCTTAACAATCAAGAGTAAATTAAGATTAACAC 1921
QY 1928 CTCAACGACTCAACAACTTTTCTGTGATTAATTAATTAATTAATTAATTAATTAATTA 1987
DB 1922 CTCAACGACTCAACAACTTTTCTGTGATTAATTAATTAATTAATTAATTAATTAATTA 1981
QY 1988 TTGGGATTTTCCAAATTTCCAAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2047
DB 1982 TTGGGATTTTCCAAATTTCCAAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2041
QY 2048 TATTTAAATCGTGAAGATGATCAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2107
DB 2042 TATTTAAATCGTGAAGATGATCAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2101
QY 2108 CAATTTAATTTCTCTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2167
DB 2102 CAATTTAATTTCTCTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2161
QY 2168 AATTAATTAATTAATTTTCCAAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2208
DB 2162 AATTAATTAATTAATTTTCCAAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2202

RESULT 3
C0868309 2085 bp DNA linear PAT 13-SEP-2004
LOCUS C0868309
DEFINITION Sequence 10 from Patent WO2004074462.
ACCESSION C0868309
VERSION C0868309.1 GI:51998355
KEYWORDS
SOURCE
ORGANISM
Bacillus thuringiensis
Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus; Bacillus
cereus group.
REFERENCE
1 Carozzi, N., Hargiss, T., Koziel, M. G., Duck, N. B. and Carr, B.
AUTHORS
TITLE Delta-endotoxin genes and methods for their use
JOURNAL Patent: WO 2004074462-A 10 02-SEP-2004;

FEATURES
source Athenix Corporation (US)
Location/Qualifiers
1..2085
/organism="Bacillus thuringiensis"
/mol_type="unassigned DNA"
/db_xref="taxon:1428"
1..2085
/note="unassigned protein product"
/codon_start=1
/transl_table=1
/protein_id="CAH33950.1"
/db_xref="GI:51998355"
/translation="MCQGNQYQDNRFPASADTTAAVSAQTVSGTLAIGGTSI
SGPIGILGALISFGTLIVFWPAGSDQKVTWQPIWGBIFVDTPLTBSIKQLKLT
LBSFROILOSYNLALDDMRKLKLOAPSSALQOALTLKIRFENVANDRIEIP
GFQLETKTLLEIYAOAAFNHLNLOQGALEADENADIDHPSQIBENAGSDDYK
LKENIPKVSNCANTYREGNLKRNENMMSIFNDYRMYTITVDTIOFSPYDIK
RYVDSIRIGIGITBELTREIYTEINFDLTVEIOPNLAIMENLITRSGELPSFID
ELIPTYKNEYGRNLVGIANKRNSYATTTETETIYGRTPPTTKTILIPESYKGLV
TDKQVPTSPFPNIYFTINQIELYLNNSPKNLTYSAGNLSNDKTTDQFPVKDC
KPIINPCLPSVNSYSHLSQFSLFVNSYKIGLALNLYTGAIGWTHS SVNRNALSD
KIITMTPAIKGNSLDNTSKVIBSGHGTGMLVYLOSGLRLEICRTFNSQSYIARL
YATNGAGTLENLSTIPGVIGIPORLNNTPSGTYNNNOYDGFQFQPSYTLPL
NRNIPPIPNRADVSNLILIDKIEFIPITSVQNRNEKXLEITIGKINLPFNHTKN
TLNIEATNIDID"

ORIGIN
Query Match 81.3%; Score 1794.4; DB 6; Length 2085;
Best Local Similarity 92.7%; Pred. No. 5.5e-234;
Matches 1908; Conservative 0; Mismatches 141; Indels 9; Gaps 2;
QY 154 ATGTGTCAAGGAATACACAAATATGATTAATTTGACACATTTGCTAGTCTGATACA 213
DB 1 ATGTGTCAAGGAATACACAAATATGATTAATTTGACACATTTGCTAGTCTGATACA 60
QY 214 ATGTGTCAAGGAATACACAAATATGATTAATTTGACACATTTGCTAGTCTGATACA 273
DB 61 ATGTGTCAAGGAATACACAAATATGATTAATTTGACACATTTGCTAGTCTGATACA 120
QY 274 CTCACTTCTAATTCGGAACGATAGAAATATAGTCTTAATTAATATCTTTTGGTACC 333
DB 121 CTCACTTCTAATTCGGAACGATAGAAATATAGTCTTAATTAATATCTTTTGGTACC 180
QY 334 CTAAATCATGCTTTTGGCCCGGGGAGAACAGACAAAACAGTAGACAAATTTAT 393
DB 181 CTAAATCATGCTTTTGGCCCGGGGAGAACAGACAAAACAGTAGACAAATTTAT 240
QY 394 AAAATGGAGAAATTTTGTGATACACGTTTAAAGAAAGCTAATAACGTTAAAGTTA 453
DB 241 AAAATGGAGAAATTTTGTGATACACGTTTAAAGAAAGCTAATAACGTTAAAGTTA 300
QY 454 CAAACTTTAAGAGATTTAGACAAATTAATTAACAAAGCTATATACAGATTAGATTGG 513
DB 301 CAAACTTTAAGAGATTTAGACAAATTAATTAACAAAGCTATATACAGATTAGATTGG 360
QY 514 AGAAATTAATAAAGACTACAAAGCTCTGATTTACACATCATCATGATTAACAAAGCT 573
DB 361 AGAAATTAATAAAGACTACAAAGCTCTGATTTACACATCATCATGATTAACAAAGCT 420
QY 574 GCCTGACCTCTTAATTAAGATTTGAGATTTGACAAATGATTTTATTCGAAATACCT 633
DB 421 GCCTGACCTCTTAATTAAGATTTGAGATTTGACAAATGATTTTATTCGAAATACCT 480
QY 634 GGTTCGAATTTAAATTTTAAACGCTATTAATTAATTTATGCGCAAGCTGCTAAT 693
DB 481 GGTTCGAATTTAAATTTTAAACGCTATTAATTTATGCGCAAGCTGCTAAT 540
QY 694 TTTCAATTAATTTATTAACAACAGGTGTAATTTGCTGATGAATGAATGACAGATATA 753
DB 541 TTTCAATTAATTTATTAACAACAGGTGTAATTTGCTGATGAATGAATGACAGATATA 600
QY 754 CATCTTCACAAATTTGAACCTAATGCTGGAACATGACATTAATTAATTTTAAAT 813

```

Db      601 CATCTTCACAAATGTAACCTAATGCTGGAACATGATGACTATTAATCTTTTAAAA 660
Qy      814 GAAATATATACCTAATATAGTAATATGTAATGCAAAATACCTATAGAACGTAATAAAAT 873
Db      661 GAAATATATACCTAATATAGTAATATGTAATGCAAAATACCTATAGAACGTAATAAAAT 720
Qy      874 CTTAGAGACGAACCAATATGAAATGAGATATATTAATGACTATCGAAGATATATGACC 933
Db      721 CTTGCAAAACGAACCTAATATGAGATGAGATATATTAATGATTAATCGAAGATATATGACT 780
Qy      934 ATTAATGATATATGATATACCACTCTCTCAATTTTCTTATATGATATATTAAGATATATGAGAT 993
Db      781 ATTAATGATATATGATATACCACTCTCTCAATTTTCTTATATGATATATTAAGATATATGAGAT 840
Qy      994 TCAATAGGAGGAATGAAATGAAATGAAATGAAATGAAATGAAATGAAATGAAATGAAATGAAAT 1053
Db      841 TCAATAGGAGGAATGAAATGAAATGAAATGAAATGAAATGAAATGAAATGAAATGAAATGAAAT 894
Qy      1054 ACTGAAATTAATTTTGAATGCTCTCTCAATTTTGAATGCTCAATTTGCTCAATTTGCTCAATTTG 1113
Db      895 ACTGAAATTAATTTTGAATGCTCTCTCAATTTTGAATGCTCAATTTGCTCAATTTGCTCAATTTG 954
Qy      1114 GAATTAATTTTGAATGCTCTCTCAATTTTGAATGCTCAATTTGCTCAATTTGCTCAATTTG 1173
Db      955 GAATTAATTTTGAATGCTCTCTCAATTTTGAATGCTCAATTTGCTCAATTTGCTCAATTTG 1014
Qy      1174 TATACAGAAATTAATTTGGAATGCTCTCTCAATTTTGAATGCTCAATTTGCTCAATTTGCTCAAT 1233
Db      1015 TATACAGAAATTAATTTGGAATGCTCTCTCAATTTTGAATGCTCAATTTGCTCAATTTGCTCAAT 1074
Qy      1234 ACTTATAGCAATTAATTTGGAATGCTCTCTCAATTTTGAATGCTCAATTTGCTCAATTTGCTCAAT 1293
Db      1075 ACTTATAGCAATTAATTTGGAATGCTCTCTCAATTTTGAATGCTCAATTTGCTCAATTTGCTCAAT 1134
Qy      1294 AAAACAAATTAATTTGGAATGCTCTCTCAATTTTGAATGCTCAATTTGCTCAATTTGCTCAAT 1353
Db      1135 AAAACAAATTAATTTGGAATGCTCTCTCAATTTTGAATGCTCAATTTGCTCAATTTGCTCAAT 1194
Qy      1354 CCGATTCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1410
Db      1195 CCGATTCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1254
Qy      1411 GGCTCATCTTAACCAACACTCAATATTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1470
Db      1255 AATTCACCTAATTAATTTGGAATGCTCTCTCAATTTTGAATGCTCAATTTGCTCAATTTGCTCAAT 1314
Qy      1471 ACAACTTTTCTCAATTTTCTAGAAATTAATTTGGAATGCTCTCTCAATTTTGAATGCTCAATTTT 1530
Db      1315 ACAACTTTTCTCAATTTTCTAGAAATTAATTTGGAATGCTCTCTCAATTTTGAATGCTCAATTTT 1374
Qy      1531 TCACCAATTTTCTCAATTTTCTAGAAATTAATTTGGAATGCTCTCTCAATTTTGAATGCTCAATTTT 1590
Db      1375 TCACCAATTTTCTCAATTTTCTAGAAATTAATTTGGAATGCTCTCTCAATTTTGAATGCTCAATTTT 1434
Qy      1591 TATGATATGATATGATATGATATGATATGATATGATATGATATGATATGATATGATATGATATGAT 1650
Db      1435 TATGATATGATATGATATGATATGATATGATATGATATGATATGATATGATATGATATGATATGAT 1494
Qy      1651 AGTGTATGATATGATATGATATGATATGATATGATATGATATGATATGATATGATATGATATGAT 1710
Db      1495 AGTGTATGATATGATATGATATGATATGATATGATATGATATGATATGATATGATATGATATGAT 1554
Qy      1711 GGTAACATTTTGAATGCTCTCTCAATTTTGAATGCTCTCTCAATTTTGAATGCTCTCTCAATTTT 1770
Db      1555 GGTAACATTTTGAATGCTCTCTCAATTTTGAATGCTCTCTCAATTTTGAATGCTCTCTCAATTTT 1614
Qy      1771 TTGGTTATTTTGAATGCTCTCTCAATTTTGAATGCTCTCTCAATTTTGAATGCTCTCTCAATTTT 1830
Db      1615 TTGGTTATTTTGAATGCTCTCTCAATTTTGAATGCTCTCTCAATTTTGAATGCTCTCTCAATTTT 1674
Qy      1831 CAATTTATTTTGAATGCTCTCTCAATTTTGAATGCTCTCTCAATTTTGAATGCTCTCTCAATTTT 1890
Db      1675 CAATTTATTTTGAATGCTCTCTCAATTTTGAATGCTCTCTCAATTTTGAATGCTCTCTCAATTTT 1734

```

```

Qy      1891 ATATCTTTAACAATACAGAGATATAGAAATACCACTCAACGACTCAACAACTTTT 1950
Db      1735 ATATCTTTAACAATACAGAGATATAGAAATACCACTCAACGACTCAACAACTTTT 1794
Qy      1951 TCTGGTACAAATTTAATTAATTTTCAATTAACGAGATTTTGGTATTTTCCATTTTCAAGT 2010
Db      1795 TCTGGTACAAATTTAATTAATTTTCAATTAACGAGATTTTGGTATTTTCCATTTTCAAGT 1854
Qy      2011 ACAGTAACATTAACCTTTAATTCGAACATACCACTTATTAATTAATTCGAGATATATCA 2070
Db      1855 ACAGTAACATTAACCTTTAATTCGAACATACCACTTATTAATTAATTCGAGATATATCA 1914
Qy      2071 AATTCATTTTAAATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2130
Db      1915 AATTCATTTTAAATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1974
Qy      2131 AATGAGAAAAACAAATTTGAAATCTATTCGAAACAAATTAATTAATTAATTTTCACAAT 2190
Db      1975 AATGAGAAAAACAAATTTGAAATCTATTCGAAACAAATTAATTAATTAATTTTCACAAT 2034
Qy      2191 CATACAAAAACACTTTGA 2208
Db      2035 CATACAAAAATTAATTTA 2052

RESULT 4
108083 108083 4451 bp DNA linear PAT 02-DEC-1994
LOCUS DEFINITION Sequence 1 from Patent EP 0296870.
ACCESSION 108083
VERSION 108083.1 GI:589204
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE Unclassified.
AUTHORS 1 (bases 1 to 4451)
TITLE Eilar,D.U. and Ward,E.S.
JOURNAL New toxin-encoding DNA fragments from Bacillus thuringiensis
FEATURES subsp. israelensis
source Patent: EP 0296870-A1 1 28-DEC-1988;
location/Qualifiers
1..4451
/organism="unknown"
/mol_type="unassigned DNA"

ORIGIN
Query Match 15.7%; Score 346.6; DB 6; Length 4451;
Best Local Similarity 53.2%; Pred. No. 5.8e-38;
Matches 1176; Conservative 0; Mismatches 934; Indels 99; Gaps 17;

Qy      4 AATCAAAATTAAGATATATGAAATATGAAATATGAAATATGAAATATGAAATATGAAATATGAAAT 63
Db      888 AATATGAAATCTTTATCAAAATTAATAATGAAATATGAAATATGAAATATGAAATATGAAATATGAAAT 947
Qy      64 CCGAAGCAAGACGATATGAAATGAAATGAAATGAAATGAAATGAAATGAAATGAAATGAAATGAAAT 123
Db      948 TTAATATATCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1007
Qy      124 CAAAACCAAAATTAACAAAGAGCTCAATATGCTCAAGGAAATATCAAAATATGATAT 183
Db      1008 CAAAGTCAAAATTTAATAAGATTGGCTCAATATGCTCAAGGAAATATCAAAATATGATAT 1067
Qy      184 AATTTGAGACATTTGCTGATGCTGATATGAAATGCTGCAATTTGCTGCAATTTGCTGCAATTTGCTG 243
Db      1068 GATTTTGAACCTTTTATGAT-----AGTGTGAACCTGAGTGTGCTTATATCTATTTGA 1118
Qy      244 TCCGTAATCTGTTAGCCGATATGATGATGCTGCTCAATTTGCTGCAATTTGCTGCAATTTGCTGCAAT 303
Db      1119 GTTGGACCGTATGATCTGCT-----TTGGGTTTCAACAACCC 1157
Qy      304 ATAGGTCTAATTAATTAATTTTGGTATACCTTAATCACTGCTTTTGGCCGCGGAGAA 363

```

Db 1158 TTAGACTTGCTTAATGAGTTTGGTACATTAATACAGTTCCTTTTCCAGCCCAAGAC 1217
Qy 364 CAAGACAAAACGATGACACAAATTTATTAATGGAGAAATTTTGTGATACACCG 423
Db 1218 CAATCTAACAC--ATGGATGATCTTATTAACAACTAAATAATATATATAAAAAGAA 1274
Qy 424 TTAAACA-GAAAGCATTAACAGCTTAAAGTACAACTTTAGAGATTTAGACAAATTT 482
Db 1275 ATAGATCAACATATATATAGTAATGCTAATAAATTTTAAACAGTCTTTATATGTTATC 1334
Qy 483 ACAAGCTATATATCAGCATATAGATGATGGAATAATTAAGAACTACAGCTCCG 542
Db 1335 AGCATTTATCTATATCACTTTAAACATGGAGAAATATCCAAACCCCAAAATATCTAG 1394
Qy 543 ATTACCAACCATCATCAGCATTAACAACAAGCTGCTTGACTCTTAATAATACATTTGAGA 602
Db 1395 GATGTAAAGG---CACAAATCCAGCTAGTTCATTACCATTTTCAAAATGTCAATCCGA 1450
Qy 603 TGTTCACAAATGATTTTATTCAGAAATACCTGTTTCCAACTTGAACTTTTAAACGCT 662
Db 1451 GCTTGTAACCTGTCTCTCTAATCTAGTGAATTC-----GATTAATAATCATACT 1504
Qy 663 ATTACCTATTTATTTAGCGAGCTGTAATTTTCATTTAATTTATTAACAACAAGTGC 722
Db 1505 AGTATTTATCTAGTATGCAACAAGCAGAACTTACATCTGATGTATTTAAATCAAGCCGT 1564
Qy 723 TGAATGGCTGATGATGAAATGACAGATATACATCTTACAAATTTGAACCTTAATGCTG 782
Db 1565 CAATTTGAAAGGATTTTAAACAAATCGAACATTCGATTTATTTAGAGCTTTGGC--- 1621
Qy 783 AACATCAGATGACTTTTAACTTTTAAAGAAATATACCTTAATATAGTAATTTG 842
Db 1622 AACGCAATGATATTAATTCACAGTATGACTAAAGCTATAGAAATTAACCTAATTTATG 1681
Qy 843 TGCAAATACCTATATGACAGGACTTAAATAATCTTAGAGACGAAACCAATATGAATGAG 902
Db 1682 TGTATCAACTTATTAACAAAGATTAATTTATTAACGACGCTGTATAGTAATCTTGA 1741
Qy 903 TATATTTAATGACTATCGAAGATATATGACATTACTGTATTTAGATCCATCTCTCAAT 962
Db 1742 TGGAAATATTAACGTGAACATTA-----CAATCGTATCGAACAAATAATGACTACTGC 1795
Qy 963 TTCTTTATATGATATATAAAGATATAGAGATTTCAATAGAGAAATAGAAATGAAGCAT 1022
Db 1796 TGTATTAATGATGTGTGCACTCTTCTCTAATTAAGATGATAGTAATTTCCAAATAGGT 1855
Qy 1023 TAAAGATGACATCAACAAGAAATTTATACAACTGAATTAATTTGATGCTTCTCTCA 1082
Db 1856 CCAATCTGAATCTTACTCGAATAATTTAT--CAAGTACTTAACCTCGAAGAAAGCCCTAT 1913
Qy 1083 ACTTAGAGTTCAACCCCAATCTAGCTACGATGAAATATATTTAAACAAGTGAAGTTTAA 1142
Db 1914 AAATATTAATGACTTTCATATATCAAGAGATTCATTAACGTAAGACC--GCATTTATTTA 1971
Qy 1143 ATTATTTTCATTTTATAGCAATTTATTTTATATACAGAAATATCAAAATTTTCGGAATCG 1202
Db 1972 CTGTGCTGATTCCTTTGAA-----TTTATAGAAAAACGCAAACTACTCTTAATTAATTT 2026
Qy 1203 TTTAGTGTATTTCTAATCGATGACCTACTTATAGCAATCTAATACTGAAGAACTTT 1262
Db 2027 TTTTACCGCACTTATATATATGTTTCAATTAACACTTATATATATATCCAAAATCTAG 2086
Qy 1263 ATATGAGAAAGAAACAGGTTCAACCAACAACAAACAAATAAGACCAATTTGAATCTTATA 1322
Db 2087 TGTTTTGGAAATCAATATGTAAGTAATTAATAATCTGTGGTTTGGCAACAAATAT 2146
Qy 1323 AGTTCAATTTGTAATGATAGCAATCACTCTGTTTCCCTATTTCAACACACTTTAT 1382
Db 2147 TTATATTTTTTTATTAATATGTCATAGCTTATATATATATTAATGATTAATATA 2206
Qy 1383 AATTAATCAAAATGAACTTTATTTAATGGCTCATTAACAACAACACTCAAAATATTCAGC 1442
Db 2207 TATATGTAATAATGATTTTTTTATTAATGTAATGTAAGTCTTTGGAGAAAGAACTTAC 2266

Qy 1443 AGGAGGCTCT-----TTATCTAATTTATCAAAACACAACCTTTTTCATTTCTAGAAA 1496
Db 2267 AGCAGATCTGGGCAATTAATCTTATGATGTAAATTAATAATTTTTCGGGTACCAATTC 2326
Qy 1497 AAAAGACTGCAATCTACTTATATGATCCAGGTGTTCACAACTTTAATACTTACTGCA 1556
Db 2327 TAAACGAAGAGAAATCAAGAAACCTTACCCTTTTCCAAATATATATTAATATGATCA 2386
Qy 1557 TATTTATCCCATTTTTCATTTATTTACTTATTTCTATGATGATGAAATTAACAGTACAAT 1616
Db 2387 TATTTTATC-----ATTTATTAAGTCTTATGTAATCCCTGCAACATAT 2429
Qy 1617 ATTAGATACAGTGTGA--TTAGATGACACACAGTATGTTGATATATATATGCAATA 1674
Db 2430 AAAATCAAGGTATATGTTTGTGCTTGACACACACTGATGTGATCTTAAATAATACAT 2489
Qy 1675 TCGATTAATAATTAATTAATGATATCCAGCAATCAAAAGTAAACATCTTGATACAACTCT 1734
Db 2490 TATACATTTTAATTAATCACTCCAAATTCAGCTGTAAAGCAATTCACATTTGGGACTGCT 2549
Qy 1735 AAGGTAATGGAAGACCTGTCATACAGAGAAACCTTGTATTTTAAACAAGTCAAGGG 1794
Db 2550 AAGGTGTTCAAGGACCTGTCATACAGAGGGAATTTATGATTTCCAAAGATCA---- 2605
Qy 1795 CGTTAGAAATTCATGATGAACTCTTAATTTTACACAAATCTTATTTCAATTTGACTTGA 1854
Db 2606 --TTCAAAATTTACATGTCAACACTCAAAATTTTCAACAATCGTATTTATTAAGAAATCGT 2663
Qy 1855 TATGCTCAAAATGATGCTGGAATAATCTCTCTAATATCTCTTACAAATACAGAGGTA 1914
Db 2664 TATGCTTCAATATGGAAGCGCAAAATCTGAGCTGTATTAATCTTATCTTATCCAGGGG-- 2721
Qy 1915 ATAGAAATACCACTCAACGACTCAACAACATTTTCTGGTATCAAAATTAATTAATTTTA 1974
Db 2722 -TAGCAAACTGGGATATAGGCACTTAACCCACTTTTCTGTATACAGTATATAGAAATTTA 2780
Qy 1975 CAATACGAGATTTTGGGTATTTCCAAATTTCCAAATGACAGTATACATTTACCTTTAAATGCA 2034
Db 2781 AAATATTAAGATTTTCACTACTTGAATTTTCTAACAGAGTGAATTTGCTCCAAATCAA 2840
Qy 2035 AACATACATTTAATTTAATTCGTCAGATGAT--CAAAATCAATTTTAATGATGAT 2091
Db 2841 AACATATCTCTGTGTTAATCGTTCGATGTATATTAACAACAACAAGTATGAT 2900
Qy 2092 AAAATGAATTTATACCAATTTACTTCTCTATGCAACCAAAATAGAGAAAAACAATAATTA 2151
Db 2901 AAAATTTGAATTTCTGCCAATTTACTCTCTTATTAAGAGATAGAGAAACAAATAATTA 2960
Qy 2152 GAAACTATCCAAACAAATAATTAATACATTTTTCACAAATCATACAAAA 2200
Db 2961 GAAACGATACAAACAATTAATTAATTAATCAATTTTATCAATCTTATAAAA 3009

RESULT 5
109103
LOCUS 109103 4451 bp DNA linear PAT 02-DEC-1994
DEFINITION Sequence 1 from Patent WO 8810305.
ACCESSION 109103
VERSION 109103.1 GI:588188
KEYWORDS
SOURCE
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 4451)
AUTHORS Ellar,D.J. and Ward,B.S.
JOURNAL Patent: WO 8810305-A 1 29-DEC-1988;
FEATURES
1. .4451
Location/Qualifiers
/organism="unknown"
/mol_type="unassigned DNA"

ORIGIN

Query Match 15.7%; Score 346.6; DB 6; Length 4451;

Best Local Similarity	53.24;	Pred. NO. 5.8e-38;
Matches 1176;	Conservative	0;
	Mismatches	934;
	Indels	99;
	Gaps	17;

4 AATCAAAATAACGATAATAACGAATATGAAATTATGATTCCATACCTCACCCTTATTTT 63

Oy	4	AAATCAAAATTAACGTAATTAACGAATATGGAATATTAATTAATGATATGCGATTAACCTCACTTAATTTT	63
Db	888	AAATATGAATCCTTTACAAATTAATAAATAATGAAATATGAAACATTAAATGATTCACAAAAAAA	947
Oy	64	CCGAACAGAAACAGTAATGATTTCTGATACCTTTCAACAAATTAATCCAAATCAACATTA	123
Db	948	TTAAATATATCTAATTAATTAATCAAGATATCCAAATAGAAAATATGTCAAAACATTAATTA	1007
Oy	124	CAAAACAAATTAACAAAGAGTGCTCAATATGTCTCAAGGAAATACAAATATGTGTAT	183
Db	1008	CAAAATTAATTAATTAAGAAATGGCTCAATATGTCTCAACAGAAATCACAGTAATGTGTGA	1067
Oy	184	AAATTCGAGATTTGCTAGTGCTGATACATATGCTGAGATTAAGCAGATTAATTTGTA	243
Db	1068	GATTTTGAACCTTATTAATTAAT-----AGTGTGAATCTCAGGCTTAATTAATTTGTA	1118
Oy	244	TCCGTAATCTGTATTAGCCGATATAGTGAGGCTCACTTCATATCCGAGCCGATAGGAATA	303
Db	1119	GTTGGGAACGTAAGTACGCTGT-----TTGCGGTTCAACAACCC	1157
Oy	304	ATAGTGCTAATTAATTAATCTTTTGTGATCCCTAATCACTGCTTTTGGCCCGGAGAA	363
Db	1158	TTAGGACTGTGTTAATTAATGATTTGGTATCAATTAATACAGATGCTTTTTCAGCCCAACAC	1217
Oy	364	CAGAACAACAGATATGAGACAATTTATTTAAATGGAGAAATTTTGTGATACACG	423
Db	1218	CAATCTAACAC---ATGAGATGACTTTATTAACAAACTAAATAATTTATTAATAAAGAA	1274
Oy	424	TTAAACA--GAAGAATMAAACAGCTAAAGTTACAACTTTAGAGGATTTAGACAAATTA	482
Db	1275	ATAGCATCAACATTAATAAGTAATGCTAATTAATAATTTTAAACAGGTGCTTTAATGTATC	1334
Oy	483	ACAAAGCTAATTAACAGATTAATGATTAATGGAGAAATTTAAATAAGCTACAAAGCTCTG	542
Db	1335	AGCACTTAATCAATTAATCACCTTAAACATGAGGAATTAATCCAAACCAACAAATTAATCAG	1394
Oy	543	ATTACACACATCAATCAGCATTAACAACAAGCTGACCTTGACTTTAAATTAATACAAATTTGAGAA	602
Db	1395	GATGTAAAGCA---CACAAATCCAGCTATGTTCAATTAACATTTTCAAATATGTCATTTCCAGA	1455
Oy	603	TGTTCAACATGATTTTAATTCGAGAAATTAATCTGTGTTTCCAACTGTAACCTTAATTAAGCGCT	662
Db	1451	GCTTGTAAACCTCTGTCTCTTAATCTTAGATGATTC-----GATTAATCTAATCAATACT	1504
Oy	663	ATTATCACTAATTTTAATGGCAAGCTGCTAATTTTCAATTTAATTAATTAATTAACAAGATGC	722
Db	1505	AGTATTAATCTAGTATGACAAAGACGAACTTAATCACTGATCTGATTAATTAATCAAGCGCT	1564
Oy	723	TGAATTTGGCTATGAAATGGAATGCGAATATACATCTTCAACAATTTGAACCTTAATGTCTG	782
Db	1565	CAAAATTTGAACGTAATTTAAAAAACAATCGACAATTCGATTAATTAAGAGCCCTTGGC---	1621
Oy	783	AAACATCAGATGACTTTTAATTAACCTTTTAAAAAATAATTAACCTTAATTAATTAATTAATG	842
Db	1622	AAACGCAATTAATTAATTAATTCAGATTAATCAATTAAGACTTAATGAAGATTAACACTTAATTAATG	1681
Oy	843	TGCAAAATACCTATATGAACAGACTTAATAAATCTTAGAGACGAACCAATATTAATGAATGAG	902
Db	1682	TGTAAACAATTAATTAATAAAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATG	1741
Oy	903	TATATTTAATGACTATCGAAGATATATGACCAATTAATCTGTATTTAGATACATCTCTCAATT	962
Db	1742	TGGAATTTAACTCGAACAACATA-----CAATCTGATCGAACAATAAATGACACTACGCTC	1795
Oy	963	TTCTTTAATATATTAATAAAGATTAATGATTAATTAATGAGGAATTAAGATTAATTAATGAGAT	1022
Db	1796	TGTATTAATATGTGTGCACTCTTCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATG	1855
Oy	1023	TAAAGAAATCACTCAAGAGAAATTTATTAACAAGTAATTAATTAATTTGATGTCTTCCCA	1082

Db	1856	CCAACTCGAACTTACTCGAGAAATTAT--CAGTACTTAACTTCGAGAAAGCCCTAT	1913
Qy	1083	ACTTAGAGTTCAACCCAACTAGCTAGATGGAATATAATTAAACAGTCAAGTTTAA	1142
Db	1914	AAATATATATGACTTCAATATCAAGAGATTCACCTTACAGTAGACC--GCATTATATTA	1971
Qy	1143	ATTATTTTCAATTTTAAACAATTTATTTTATACGAAATATCAAAATTTGGGAATCG	1202
Db	1972	CTTGCGCTGATCTTTGAA-----TTTTATGAAAAAGCCAAACCTCTCTATATATTT	2026
Qy	1203	TTTAGTTGGTATTTCTATCGTATGCACTTACTTATAGCAATATCTATTACTGAACTTT	1262
Db	2027	TTTACACGACCATTTATATATGTTTCATTACACCTTGATATATATATCCAAAATCTAG	2086
Qy	1263	ATATGGAAGAAAGAACAGGTCACCCACAACAAAACAAATTAAGCATTGAACTTTAA	1322
Db	2087	TGTTTTGGAATACACATGTATCTGATTAATTAATTAATCTCTGGTTTGGCACAAATAT	2146
Qy	1323	AGTTTCAATTGTATCTGATAGACATCAGCTCTGTTTCCCTATTTCACACAATTAT	1382
Db	2147	TTATATTTTTTATTAATATGTCAATAGCTTAGATATATTAATATCTAAATGATTATATA	2206
Qy	1383	AATTATCAAAATGAACTTTATTTAAATGCTCATCTTACACACACTCAATATTTGAC	1442
Db	2207	TATTAGTAAATGATTTTATTTTATATACATATGATCTAGACTTTTGGAGAAAGAACTTAC	2266
Qy	1443	AGGAGGGCT-----TTATCTATAATTACAAAACACAACCTTTTTCATTTCCCTAGAA	1496
Db	2267	AGCAGAGATCGGGCAAAATTAATTAATGATATTAATTAATTAATTTTCGGTTTACCAATCT	2326
Qy	1497	AAAGACTGCAATCTAGTTATTTGATCCAGGTGTTCCACAACTTTAATTAATATAGTCA	1556
Db	2327	TAAACGAAGAGAAATCAGAGAAACCTACCCCTTTTCCAAACATATGATTAACATATGCA	2386
Qy	1557	TATTTTATCCCATTTTCAATTTATTACTTATCTATATGATGATGGAATTACGCTACAAT	1616
Db	2387	TATTTTATC-----ATTATTAATAAGTCTTATGATCCCTGCAACATAT	2429
Qy	1617	ATTGATATACAGTGTGA--TTAGATGACACACAGTGTGTGATGATATTAATGCAATA	1674
Db	2430	AAAATCAGTATATAGTTAGTTGCTTGGACACACTATATGTATGATCTTAATAATACATTT	2489
Qy	1675	TCAGATAAATTAATTAACAATGATCCAGCAATCAAGATTAACAATCTGTGATACAACTCT	1734
Db	2489	TATACACTTTTAACTACCCAAATTCAGCTGTATTAAGAAATTCACCTTGGGACGTCTCT	2549
Qy	1735	AAGGTAAATGAAGACCTGTCATACAGAGGAACTTGGTTATTTTCAAAGTCAAGGG	1794
Db	2550	AAGGTGTTCAGAGACCTGTCATACAGAGGGGATTTAATTGATTTCAAAAGATCA----	2605
Qy	1795	CGTTTAAATTAATATGATGAAACCTCCTAATTTCTACAAATCTTATTTTCAATAGACTCGA	1854
Db	2606	--TTTCAAAATTAATGTATCAACTCAAAATTTTCAACAAATCGTATTTTATTAAGAAATCGT	2663
Qy	1855	TATGCTCAAAATGSGTGTGAAATACCTTCTATATATCTCTTACAAATACAGAGTAT	1914
Db	2664	TATGCTTCAATGGAAGCGAAATATCTCGAGCTGTATTAATCTTAGTATCCACAGGG--	2721
Qy	1915	ATAGGAATACCACTCAACGACTCAACAACAATTTTCTGGTCAAAATTTATTAATATTTA	1974
Db	2722	-TAGCAGAACTGGGTATAGCACTCAACCCACATTTTCTGTATCAGATTTATACGAATTTA	2780
Qy	1975	CAATACGGAATTTTGGGTATTTCCAAATTTTCCAAAGTACATTAACATTAACCTTTAAATCGA	2034
Db	2781	AAATATTAAGATTTTCACTACTTGAATTTTCTTACAGAGGTGAATTTGCTCCAAATCAA	2840
Qy	2035	AACATACCATTTATTTAAATCGGACAGATGAT--CAATTCATTTTATTAATCAATGAT	2091
Db	2841	AAATATCTCTGTGTTTATATCGTGTGAGATGATATATCAACAACAACAGTACTTATGAT	2900
Qy	2092	AAAAATGAATTTATACCAATTAATCTCTCTATGCAACAAAATAGAGAAAACAAAATTA	2151

Db 2901 AAAATTGAATTCTGCCATTACTCTGTTCTATTAGAGAGGATAGAGCAAAATTTA 2960

Qy 2152 GAAACTATCCAAACAAATTAATACATTTTTCACAAATCATACAAAA 2200
|||||
Db 2961 GAAACAGTACAAACAAATTAATACATTTTATGCAAAATCCTATAAAA 3009

RESULT 6

BTX0D1

LOCUS BTX0D1 3543 bp DNA linear BCT 18-APR-2005

DEFINITION *Bacillus thuringiensis* gene for 130 kDa delta-endotoxin.

ACCESSION Y00423

VERSION Y00423.1 GI:40351

KEYWORDS delta-endotoxin; endotoxin.

SOURCE *Bacillus thuringiensis*

ORGANISM *Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus; Bacillus cereus* group.

REFERENCE 1 (bases 1 to 3543)
Ward, E.S. and Ellar, D.J.
Nucleotide sequence of a *Bacillus thuringiensis* var. *israelensis* gene encoding a 130 kDa delta-endotoxin
Nucleic Acids Res. 15 (17), 7195 (1987)

JOURNAL PUBMED 2821500

REFERENCE 2 (bases 1 to 3543)
Ward, E.S.
Direct Submission
Submitted (11-MAY-1987) Ward E.S., University of Cambridge,
Department of Biochemistry, University of Cambridge, Tennis Court
Rd., Cambridge CB2 1QW
*strain= var. *israelensis*;
Data kindly reviewed (12-JAN-1988) by Ward E.S.

COMMENT FEATURES

source 1. .3543
/organism= "*Bacillus thuringiensis*"
/mol_type= "Genomic DNA"
/db_xref= "taxon:1428"
1. .3543
/note= "unnamed protein product; endotoxin (AA 1-1180)"
/codon_start= 1
/transl_table= 1
/protein_id= "CA66485.1"
/db_xref= "GI:40352"
/db_xref= "GOA:P16480"
/db_xref= "UniProt/Swiss-Prot:P16480"
/translation= "MNPYQNRKREYETLAAQSKLNISNNYTRYPEIENSPKQLOSTNY
KDMINMCOQOYGGDFETIDSGELASVTIVGVLGFGFTPLGLALIGFTLLP
VLPAODSNTMSDPIOTKTIKKELASITVSNANKLINSFVISTYHNHKTWEN
NPNQNTDVRTOQLVHYHFQNYIPELVNSCPNPSCDYNTLVLSYAKANLILK
TVLQAVFEKAYLKNRQFDYLEPPLAIDYPAVTAIBDYNYCVTTAKGLNLK
TTPSNLNGNINMNTYNTYRTKMTAVADVALLPNDVGRYPGVQSELRELYQVL
NFESPYKYDFOYQOESLTRPHLFTLDELNFEKACQTPNPFISHVMFHYTLN
NISQSVFQGNHVTDLKSLGATNIYIFLNLVSLDNKYLNNYSKMDFTING
TRLLEKELTAGSGOITVYDNKNIKGLPIKREKQNGPTLPPTDNTSHLSFKISL
IPATKYQVYTFPATHSVDPKNTITHTLTQIPAYANSLGTASRYVQGGHGTGL
IDFDHFKITQSHNFQSYIRIRIVANSANRAVAVNISIPVAEKGALNFTSG
TDYNLKXKQDYLEPSPNEKFPAPQNISLVFNRSDVYTTVLIIDIKIELPITRSR
EDREKQKLETVQOILNFEYANPIKNTLOSLELTDYDIDQAAVLVCSISELTPKMRLL
LDDEVNAKOLSGSNVLONGDFESATLGMTSDNITIOEDPIFKGHYHSGARDID
GTFEPYVIFOKSIDSKLKPYTRIVRGVGSKQVELVLSVGESEIDAIMNVPADLWY
LYPSTPCGSGNRCEBSAVPANIGTSMILSCQDYTKKHYVCQDSHQSFTIDTA
LDITENIGVMWFKISSPDYASLDNLEVEISGPIDGALSRYKHEKKNWDMQEAR
SETQOAVDAKQALDALFTNVQDSLEQDITLQIQVAEYLVSIPYVNDMLSDVRS
MNYDIYVELDARVAQARYLIDIRNIKNGDFTQGMVHVGANADVQOIIDVSTLVLS
NMSAGVSONVHLQHNHGVILGVIAKSGPGNGVYTLMDWESNQEKLFTSCSESYITK
TVDPPTDTRIRIEIGTBESFYBSEILQME"

old_sequence 919
/note= "g was a in [1]"
922
/citation= [2]
old_sequence 922
/note= "g was a in [1]"
/citation= [2]
ORIGIN

Query Match 15.6%; Score 345.4; DB 1; Length 3543;
Best Local Similarity 53.2%; Pred. No. 9.2e-38;
Matches 1173; Conservative 0; Mismatches 931; Indels 99; Gaps 17;

10 AATPAGATTAATAAGCAATTAATGAATTAATGATTGCGATCCCTCACTTATTTTCCGAC 69
|||||
Db 4 AATCTTATCAAAATTAATAAATGAATATGAACATTAATATCTTCACAAAATTTAAAT 63

Qy 70 AGAAGCAGTATGATTTCTAGATACCTTACACAAATTAATCAATCAATCAATCAAAAC 129
|||||
Db 64 ATATCTAATTAATTAATCAAGATATCCAAATGAATAATGTCCAAAACAAATTAATCAAAAGT 123

Qy 130 ACAATTAACAAGAGTGGCTCATATATGTGCAAGGGAATACAAATATGTGATTAATTC 189
|||||
Db 124 ACAATTAATAAGATGCGCTCAATATGTGCAACAGATGACGATATGTGTGAGATTTT 183

Qy 190 GAGACATTTGCTAGTGTATACAAATTCGTCAGTTAGTGCAGGTAATTTGTATCCGGT 249
|||||
Db 184 GAAACTTTATTTGAT-----AGTGTGAACCTAGTCTATTAATTTGTTGGG 234

Qy 250 ACTCTGTTAGCCGGTATAGGTGGGCTCATCTTATATCCGACCGATAGGAATTAATAGT 309
|||||
Db 235 ACCGTACTGACTGGT-----TTGGGGTTCAACAACCCCTTGA 273

Qy 310 GCTAATAATATATCTTTTGTGACCTTATATCATGTCTTTTGGCCCGGGAACAAGAC 369
|||||
Db 274 CTGCTTTAATATAGGTTTGTGATTAATATCAAGTCTTTTCCAGCCCAAGCAATCT 333

Qy 370 AAAACGATATGACACAAATTTAATAATGGAATTTTGTGTATACCGGTATACA 429
|||||
Db 334 AACCC--ATGAGTGAACCTTTATTAACCAAACTAATAATTTTATTAATAAAGAAATAGCA 390

Qy 430 -GAAAGCATTAACACGCTAATAGTTACAACTTTAGAGATTTAGACAAATATTTACAAG 488
|||||
Db 391 TCAACATATATATGATATGCTAATTAATAATTTTAAACAGTGTATATATGATACGACAT 450

Qy 489 CTATATATACGACTTATGATATGAGAAATTTAAAGACTACAGCTCTGATTAAC 548
|||||
Db 451 TATCATATATCACTTAAATGAGAAATATCCAAATCCACAAATATCTACAGATGTA 510

Qy 549 ACCATATACGACTTAAACAGAGCTGCTGACCTTAAATATGAGATTGAGAAATGTTCA 608
|||||
Db 511 AGGA---CACAATTCAGCTAGTTACTTACATTTTCAAAATGTCACTTCCAGAGCTTGT 566

Qy 609 CAATGATTTATTTGAGAAATACCTGTGTTTCAACTGAACTTAATAACGCTATTAAT 668
|||||
Db 567 AAATCTTTGCTCTTAATCCTAGTATGCG-----GATTACTATTAACATGATGATAT 620

Qy 669 AACTATTATGCGCAGCTGCTAATTTTCAATTAATTTATTAACAACAGTGTGAAAT 728
|||||
Db 621 ATCTAGTATAGCAACACACAAACTTACATGCACTGATTAATCAACGCCGCAAT 680

Qy 729 GCGTATGAAATGGAATGCAAGTATACATCCCTCAACAATGAACCTAATGCTGGAATC 788
|||||
Db 681 TGAAGCTATTTTAAAAACAATGCAATTCGATTAATTTAGAGCTTTGCT--AACAC 737

Qy 789 AGATGACTATTAATACTTTTAAAAAGAAATATATCTTAATATAGTAATATTTGCAAA 848
|||||
Db 738 AATTGATTAATTAATCAATATTAAGCTAATAGAGATTAAGCTAATTAATTTATGTTGAC 797

Qy 849 TACTATATGAAACAGACTAATAAACTTTAGACGAAACCAATATGAAATGAGATATTT 908
|||||
Db 798 AACTTATTAATAAGATTAATTAATTAATAAAGCAACCCGATGATATCTTGTAGAAA 857

Qy 909 TAAATGACTATGAAATATATGACATTAATGATTAATGATACCATCTCTCAATTTTCTT 968
|||||
Db 858 TATAAATCGAACAATA-----CAATAGCTATGAAACAAAATATGACTGCTGATTT 911

Qy 969 ATATGATATATAAAGATATAGATTCATATAGAGAAATAGAGAAATAGGATTAAGAA 1028
|||||
Db 912 AGATGTTTGGCACTCTTCTTAATTAATGATGTAATGTAATATCAATATAGGTGTCATC 971

1029 TGAATCTACAGAGAAATTTATACAACTGAAATTAATTTGATCGTCTTCAACTAG 1088
1089 AGTTAACCCATCTAGCTACGATGATATATTTAACAGCTGCAAGTTTAAATAT 1148
1030 TATGACTTTCATATCAAGAGATTCACCTTACAGCTAGACC--GCATTTATTTACCTGGC 1087
1149 TCAATTTTACAGCAATTTATTTTATATACAGAAATTAATTTGCGGAATCGTTAGT 1208
1088 TTGATCTTTTGA-----TTTATATGAAAGCGAACTACTCTCAATATATTTTTCAC 1142
1209 TGGTATTTCTATCCGATGACCTACTATATAGCAATCTATACCTGAACTTTATATG 1268
1143 CAGCAATATATATATGTTTCACTTACACCTGATATATATATCCAAAATCTAGCTTTT 1202
1269 AGAAAGAAAGGTTCAACCCACAAACAAATAGAACCTTTGAATCTTATTAAGTTTC 1328
1203 TGGAAATCAATGATGATGATTAATTAATCTCTGGTTGGCAACAAATATTTATAT 1262
1329 AATTGTAATGATAGCAATCACTCCCTGTTTCCCTATATGACACACTTTATATATTA 1388
1263 TTTTATATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1322
1389 TCAATTTGAACTTTATTTAATGCTCATCTAACAAACACTCAATATTTTACAGCAGAG 1448
1333 TAAATGATTTTTTTTATATACATAGTACTGATCTTTGGAGAAAGAACTTACGACAG 1382
1449 GTCT-----TATCTAATTTATCAAAACAACTTTTTCATTTCTTGAAGAAAGAA 1502
1383 ATCTGGCAAAATTAATCTTATGATGATGATTAATTAATTTTTCGGGTTTACCAATCTTAAACG 1442
1503 CTGCAATCTAGTATTTATGATCCAGCTGTTCCAGCACTTTAATATCTATGATGATTTT 1562
1443 AAGAGAAATCAAGAAACCTTACCTTTTTCACATATGATGATGATGATGATGATGAT 1502
1563 ATCCATTTTCACTATTTTACTTATCTATGATGATGATGATGATGATGATGATGATG 1622
1503 ATC-----ATTATTAAGTCTTATGATGATGATGATGATGATGATGATGATGATG 1545
1623 TACAGCTGTA--TTAGATGACACACAGTATGTTGATGATGATGATGATGATGATGAT 1680
1546 CAAGGTATAGCTGTTGTTGACACACTGATGATGATGATGATGATGATGATGATGAT 1605
1601 AAAATTAATCAATATCCAGCACTCAAGGTAATCTTATGATGATGATGATGATGATG 1740
1606 CATTTAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAAT 1665
1741 ATTGAAGACCTGATCATACAGAGGAACTGTTTATTTTACAAAGTCAAGGCTTTA 1800
1666 GTTCAAGACCTGATCATACAGAGGAACTGTTTATTTTATTTTCAAGATCA-----TTTC 1719
1801 GAAATTAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAAT 1860
1720 AAAATTAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAAT 1779
1861 ACAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1920
1780 TCAATTAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAAT 1836
1921 ATACACCTCAACGATCAACCAACTTTTGTGATCAAAATTAATTAATTAATTAATTAAT 1980
1837 GAACTGGGTAATGATCACTCAACCCACTTTTGTGATCAATTAATTAATTAATTAATTAAT 1896
1981 GGAATTTTGGTATTTTCAATTTTCAAGTACAGTACATTAATCTTAAATGAAACATA 2040
1897 AAAATTTTCACTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1956
2041 CCAATTAATTAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAAT 2097
1957 TCTCTTGTGTTTAACTGTTGATGATGATGATGATGATGATGATGATGATGATGATGAT 2016
2098 GAATTTATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAAT 2157

Db 2017 GAAATTTGCGCAATTAATCTGTTCTATACAGAGATGATGAGAAACAAATTAAGAAC 2076
Qy 2158 ATCCAAACAAATTAATCAATTTTTCACAAATCAATCAATCAATCAATCAATCAAT 2200
Db 2077 GTACAAACAAATTAATCAATTTTATGCAAAATCTTATTAATAA 2119

RESULT 7
BACISR4
LOCUS
DEFINITION
BACISR4
4253 bp DNA linear BCT 17-FEB-1998
Bacillus thuringiensis israelensis plasmid gene for 130 kDa
insecticidal protein (ISR4), complete cds.
D00248
D00248.1 GI:216289
130 kDa insecticidal protein (ISR4).
Bacillus thuringiensis serovar israelensis
Bacillus thuringiensis serovar israelensis
Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus; Bacillus
cereus group.
1 (bases 1 to 4253)
Sen, K., Honda, G., Koyama, N., Nishida, M., Neki, A., Sakai, H.,
Himeno, M. and Komano, T.
Cloning and nucleotide sequences of the two 130 kDa insecticidal
protein genes of Bacillus thuringiensis var. israelensis
Agric. Biol. Chem. 52, 873-878 (1988)
The genes of ISR3 and ISR4 that were 130 kDa insecticidal
proteins of Bti were sequenced and compared with other insecticidal
protein genes. ISR3 and ISR4 were identical in a region of the
C-terminal 467 amino acids.
Location/Qualifiers
1..4253
/organism="Bacillus thuringiensis serovar israelensis"
/mol_type="genomic DNA"
/strain="HD522"
/sub_species="israelensis"
/db_xref="taxon:1430"
/clone="pBGH4"
/plasmid="pBT1-5 and pBT1-6"
/note="976 bp downstream of HindIII site"
381..385
/note="Shine-Dalgarno sequence"
393..3935
/codon_start=1
/transl_table=11
/product="130 kDa insecticidal protein (ISR4)"
/protein_id="BA00179.1"
/translation="MNPKONKEVETLNASQKLNISNNRYPIENSPOKLOSTNY
KDMNMOONQVGGDEPFTDGLSLASVTVGTVLNGPPTPLGLALGFTLIP
VLPRAQDSNTMSDFITOTKRIITKELASTYISNANKLINSFVSYTHNHLKWMN
NPNPONTODVARTQIOLVHYHFOVNIPELVNSCPSPSCDYNILVLSYAOANLHL
TVLAQVAFKAVYLKNNRQFDLBPALIDYVPVATKIDBYNYCVTTYKGLNLIK
TTPDSNLGNNINNTYVYTKMTAVLIDVLPNVDGKVPVGOSELTRBYOVL
NFESPYKYVDFORQEDSLTRPHLFTWLDLNFKEKOTCPNNFETSHYMFHTLD
NISOKSPVGNHNTLDLKSGLATNFIPLNYLILNKLINDYNNISKDPFTLNG
TRLLEKELTAGSGQITDYVKNITFGPLIKRBNQNTLPPTYDNYSHLISFKLS
IPATIKQVYTFPANTSSVDPKNTITLHTQIIPAVKANSGLNASKVQGFHTGDL
IDPDHFETIKQHSNFOQSPYRIRVANSNGANRAVNTNISIPGVALGMLNPTFSG
TDYTNLKKDFQVLEFSPSNEVKEFAPNONSIVFNRSVDVYNTVTLIDKLEPLTSSIR
EDREKOKLETVOQIINFYANPIKNTLOSLELTDIDDAANLVECSISELYPKEMLL
LDDEVNKKLOSRSNVLONDPRESATIGMTSDNITDQEDPIPKGHVLAHSGARDID
GTIPPTIYFOKIDSKLKPRTYRYLVKGVSSKQVLEYVSYGEISDIAIMVPADNY
LYSTFDEGSNRCBTSAPVANIQTNSDMLISQYDTGKRVQCDSSHQFSTIDTGA
LDITENIGVWVWFKISSPDYASLDNLEVEEGPIDGALSRYKMEKMDQMAKRS
SETQOADVAKQALIDALFTNVQDALQDFTLADIQVAYEVVOSIPYVYNMQLSVPG
MNYDIYELDARVQARLYLDTBNIKNGDFTQGVGMHVNQMDVQOIDEVSVLVS
NMSAGVSONVHLOHNGVYLVRIAKKSPGKGYTLMDCENQKRLFTSCBEVYTK
TVDVFPPTDRIRIIGRTESFYIESILICMNE"

ORIGIN
Query Match 15.6%; Score 345; DB 1; Length 4253;
Best Local Similarity 53.2%; Pred. No. 9.7e-38;

Matches 1175; Conservative 0; Mismatches 935; Indels 99; Gaps 17;
QY 4 AATCAAAATACGATATATACGAATATGAAATTATGATTCGCATACCTCACTTATTT 63
Db AATATGAAATCCCTATCAAAATAAAAATGATATGAAACATTAATATGCTTCACAAAAAAA 449
QY 64 CCGAAGCAAAACAGTATGATTTCTGATACCTTACACAAATTAATCCAAATCAACATTA 123
Db 450 TTTAATATATCTAATTAATTAATACAGATATCAATAGAAAATAGCCAAAACAAATTATTA 509
QY 124 CAAAACACAAATTCAAAGAGTGGCTCAATATGTCGAAAGGAAATACAAATATGTTGAT 183
Db 510 CAAAGTCAAAATTAATGAAGTTGGCTCAATATGTCGAAAGGAAATACAGATATGTTGAT 569
QY 184 AATTCGAGACATTTGCTAGTGTGATACAAATGCTGACAGTTAGTGCAGGATCTATTTGA 243
Db 570 GATTTGAAATCTTTATATGAT-----AGTGTGAATCTCAGTGCCTATATGTTGTA 620
QY 244 TCCGGTACTCTGTTAGCCGGTATAGGTGGGCTCACTTATATCGGACCGATAGAAATA 303
Db 621 GTTGGACCGTACGACTGT-----TTCCGGTTCACACACACC 659
QY 304 ATAGGTCTATATATATCTTTTGGTACCCCAATCACTGTCTTTTGGCCGGGAGAA 363
Db 660 TTAGGACTTGGCTTATATAGTTTGGTACATTAATACAGATTTCTTTTCCAGCCCAAGAC 719
QY 364 CAAACAAACAGTATGACACAAATTTATTAATGGAAGAAATTTTGTGATACACG 423
Db 720 CAATCTAACAC--ATGAGTGAATTTATTAACAAACTAATAAATATTAATAAAAAGAA 776
QY 424 TTAACA-GAAAGCATTAACAGCTAAAGTACAACTTTAGAGATTTTACAAATATT 482
Db 777 ATACATCAACATATATAGTATGCTAATAAATTTTAAACAGTCCCTTAAATGTTATTC 836
QY 483 ACAAGCATATATACAGCATTAGATGATGGAAGAAATTAAGAACTACACAGCTCCG 542
Db 837 AGCATTAATCAATACCTTAAACATGAGAAATTAATCCAAACCCACAAATATCTAC 896
QY 543 ATTACCAACCATCATCAGCATTAACAACAAGCTGCTTGACTCTTAATATACATTTGAGA 602
Db 897 GATGTAAAGG----CACAAATCCAGCTAGTTCAATTACATTTTCAAAATGTCAATTCGAA 952
QY 603 TGTTCACATGATTTTATTCGAGAAATACCTGTTTCCAACTTGAACCTTAAAGCT 662
Db 953 GCTTGAATACCTGTCTCTCTTAATCTTAGATGATTC-----GATTAATCTTAACATATCT 1006
QY 663 ATTACTACCTATTTATGCGAAGCTGCTAATTTTCAATTAATTAATTTATTAACAAGGTC 722
Db 1007 AGTATTAATCTAGTTATGCAACAGCAGCAACTTACATCTGATTAATTAATCAAGCGCT 1066
QY 723 TGAATGGCTGATGAATGGAATGCAGATATACATCTTCACAAAATGGAACCTTAATGCTGG 782
Db 1067 CAAATTTGAACGATTTTAAAAAACAATCGAACATTTGATTTTAGAGCTTTGGC--- 1123
QY 783 AACATCAGATGACTTATTAATCTTTTAAAAAGAAATATACCTAATATTAAGTAACTATTG 842
Db 1124 AACGCAATTAATTAATTAATCAAGTATGACTTAAGCTATAGAAATTAACCTAATTTATG 1183
QY 843 TGCAAATACCTATATGAACAGGACTTAAAAATCTTAGAGACGAAACCAATATGAATGAG 902
Db 1184 TGTAAACAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1243
QY 903 TATAATTAATGACTATCGAAGATATATGACCATTAATCTGATTAAGATACCATCTCTCAAT 962
Db 1244 TGAATAATTAATCGAAGACATA-----CAATACGATGGAACAAAAATGACTACGCG 1297
QY 963 TTTCTTATATGATATAAAAAGATATAGATTTCAATAGAGGAATAGAAATGAAAGCAT 1022
Db 1298 TGTATTTGATCTGTGTGACTCTTCTTAATTAATGATGATAGTAAATATCAAAATAGGTGT 1357
QY 1023 TAAAGATGAATCAACAAGAAATTTATACAACTGAAATTAATTTGATGCTTCTTCA 1082
Db 1358 CCAATCTGAATCTACTCGAATAATTTAT--CAGGTACTTAATCTCGAAGAAAGCCCTAT 1415

QY 1083 ACTTAGAGTTCAACCCATCTAGCTACATGAGATTAATTAATTAACAGCTGCAAGTTTAA 1142
Db 1416 AAATATATGACTCTTCAATATCAAGAGATTCATTCACAGTAGACC--GCATTTATTTA 1473
QY 1143 ATTAATTTCAATTTTAAACAATTTATTTTATATACGAAAAATATCAATTTTGGGAATCG 1202
Db 1474 CTGGCTTGATCTTTGAA-----TTTTATGAAAAAGCCCAAACTCTCTAATTAATTT 1528
QY 1203 TTTAGTTGATTTTCTAATCGTATGACACCTTATAGCAATATCTAATTAATCTGAACCTTT 1262
Db 1529 TTTCAACAGCATTAATTAATATGTTTCAATACAGCTTGATTAATATATCCAAAAATCTAG 1588
QY 1263 ATATGGAAGAAAGACAGTTTCAACCCACACAAAAACAATAAGCAATTTGAATCTTAATA 1322
Db 1589 TGTTTTGGAAATCACAATGTAACTGATTAATTAATTAATCTTGGTTGGCAACAAATAT 1648
QY 1323 AGTTCAATTTGATAGTATAGCAATACCTCTGTTTCCCTATTCACCAACCTTAT 1382
Db 1649 TTAATATTTTATTAATATGATAGCTTAGATTAATTAATTAATTAATTAATTAATA 1708
QY 1383 AATTAATCAATTAATGAACTTATTTAAATGCTCATCTAACACACACTCAAAATATTCAGC 1442
Db 1709 TATTAATGAATGATTTTATTAATTAATTAATGATAGTACTAGACTTTGGAGAAAGAACTTAC 1768
QY 1443 AGAGGGTCT-----TTATCTAATTAATCAAAACAACTTTTTCATTTCTAGAAA 1496
Db 1769 AGCAGATCTGGGCAATTAATTAATTAATTAATTAATTAATTTTCCGGTTACCAATTCCT 1828
QY 1497 AAAAGACTGCAATCTAGTATATGATCCAGGTGTTCACAACTTTAATAATTAATTAATCA 1556
Db 1829 TAAACGAAGAGAAATCAAGAAAACCTTACCTTTTTCACAAATATGATTAATTAATCA 1888
QY 1557 TATTTATCCATTTTCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1616
Db 1889 TATTTATCT-----ATTATTAATTAATCTTAGTATCCCTGCAACATAT 1931
QY 1617 ATTAGATACAGGTGA--TTAGATGGAACAACAGTATGTTGATATGATTAATGCAATA 1674
Db 1932 AAAATCTAAGTATATAGTTTGGCTTGGACACACTGTAGTGTGATCTTAAATAATACATTT 1991
QY 1675 TCGATTAATAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1754
Db 1992 TATACATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2051
QY 1735 AAGTAATTAAGAGCTGCTGATACAGAGAACTTGGTTATTTTAAAGTCAAGG 1794
Db 2052 AAGTTGTTCAAGGACTGCTGATACAGAGGGAATTAATTAATTAATTAATTAATTAATTA 2107
QY 1795 CGTTTGAATAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1854
Db 2108 --TTTCAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2165
QY 1855 TATGCTCAAAATGCTGGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1914
Db 2166 TATGCTCAAAATGGAAGGCAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2223
QY 1915 ATAGGAATTAACACTCAAGCACTCAACAACTTTTCTGGTCAAAATTAATTAATTAATTA 1974
Db 2224 -TAGCAAACTGGGATATAGGACTCAACCCACTTTTCTGGTCAAAATTAATTAATTAATTA 2282
QY 1975 CAATACGAGATTTTGGGATTTTCAATTTTCAAGTACAGTAACTTAATTAATTAATTAAT 2034
Db 2283 AAATATTAAGATTTTCAATTTTCAATTTTCAAGGATTAATTAATTAATTAATTAATTAAT 2342
QY 2035 AACATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2091
Db 2343 AACATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2402
QY 2092 AAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2151
Db 2403 AAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2462

QY	1443	AGAGGGGCT-----TATCTAATTATCAAAACAACTTTTTCATTTTCTAGAAA	1496
Db	2745	AGCAGATCTGGGCAATTACTTATGATGTAAATAAAATATTTTCGGGTACCAATTC	2804
QY	1497	AAAAGCTGCATCTAGTATATGATCCAGGTGTCACAACTTAAATCTATAGTCA	1556
Db	2805	TAAACGAAGAGATCAAGGAAACCCATCCCTTTTCCATCATATGATTACTATAGTCA	2864
QY	1557	TATTTTATCCCATTTTTCATTATTTACTTATTTCTATGTGATTTGGATTACAGCTCAAT	1616
Db	2865	TATTTTATC-----ATTATATMAAGCTTATGATATCCCTGCAGCAATAT	2907
QY	1617	ATTAGATTCAGGTGTA--TTAGATGACACACAGTAGTGTGATATATATGACATA	1674
Db	2908	AAAATCTAAGGTATACGTTTGTCTTGACACACTCTAGTGTGATTCCTAAATAATAT	2967
QY	1675	TCAGATAAATTAATTAACAATGATCCAGCAATCAAAAGGTAACTTTGATACAACTCT	1734
Db	2968	TATACACTTAACTAACCCAAATTCAGCTGTAAAGAGAAATCACTTGGGACGTCTCT	3027
QY	1735	AAGTAAATGAAGACCTGCTCATACAGAGAAACCTTGTTTATTTACAAAGTCAAGGG	1794
Db	3028	AAGGTGTGTCAGAGACCTGCTCATACAGAGAGGGATTTAATTTGATTTCAAAAGATCA---	3083
QY	1795	CGTTTAGAAATTACATGTGAAACCTCTAATTTCTACACAACTCTTATTTCAATAGACTTGA	1854
Db	3084	--TTTCAAAATTTACATGTGCAACTCAAAATTTTCAACATGTATTTATTAAGAAATTCGT	3141
QY	1855	TATCTCAAAATGAGTGTGCAAAATACCTCTCTAATATATCTCTTACAAATCCAGAGATA	1914
Db	3142	TATGCTTCAATATGAAAGGCAAAATATCTGACCTGTATTAATCTTAGATCCAGGGG--	3199
QY	1915	ATAGGAATACCACTCAACGACTCAACACACTTTTCTGGTACAAATTAATTAATTTA	1974
Db	3200	-TACAGAACTGGGATATGACACTCAACCCACTTTTCTGGTACAGATTAACAAATTTA	3258
QY	1975	CAATACGAGAAATTTGGGTATTTCCAAATTCGAAGTACAGTAACATTAACCTTTAAATCGA	2034
Db	3259	AAATATAAAGATTTTTCAGTACTTGAATTTCTTAAGAGGAAATTTGCTCCAAATCA	3318
QY	2035	AACATACATTTATTTAATTCGAGCAATGTAT---CAATTCAAATTTTATCATTTGAT	2091
Db	3319	AACATATCTCTTGTTGTTAATCTGTTGGATGTATATACAAACACACAGTACTTATTTGAT	3378
QY	2092	AAAAATGAATTTATACCAATTACTTCTCTATGACCACAAAATATAGAGAAAACAAAATTA	2151
Db	3379	AAAAATGAATTTCTGCGCAATTACTCGTTCTATTAAGAGAGATAGAGAAACAAAAATTA	3438
QY	2152	GAAACTATCCAAACAAAAATTAATATCATTTTTCACAAATCATACAAAAA	2200
Db	3439	GAAACAGTACAAACAATTAATTAATACATTTTATGCAAAATCCTATAAAAA	3487
RESULT 9			
BPTPTOXIS/c			
LOCUS	BTPTPTOXIS	127923 bp	DNA linear BCT 16-APR-2005
DEFINITION	Bacillus thuringiensis subsp. israelensis plasmid pBctoxis.		
ACCESSION	AL731825		
VERSION	AL731825.1	GI:21685410	
KEYWORDS	Bacillus thuringiensis serovar israelensis Bacillus thuringiensis serovar israelensis Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus; Bacillus cereus group.		
SOURCE	1 (bases 1 to 127923)		
ORGANISM	Bacillus thuringiensis serovar israelensis, Murphy, L., Quail, M.A., Holden, M.T., Harris, D., Zaritsky, A. and Parkhill, J. Complete sequence and organization of pBctoxis, the toxin-coding plasmid of Bacillus thuringiensis subsp. israelensis Appl. Environ. Microbiol. 68 (10), 5082-5095 (2002)		
REFERENCE	2 (bases 1 to 127923)		
JOURNAL	Parkhill, J.		
PUBMED	12324359		
REFERENCE	2 (bases 1 to 127923)		
AUTHORS	Parkhill, J.		

FEATURES	source
COMMENT	<p>Direct Submission</p> <p>Submitted (19-APR-2002) Submitted on behalf of the pBtoxis sequencing team, Sanger Centre, Wellcome Trust Genome Campus, Hinxton, Cambridge CB10 1SA E-mail: parkhill@sanger.ac.uk</p> <p>Notes:</p> <p>Details of pBtoxis sequencing at the Sanger Centre are available on the World Wide Web.</p> <p>(URL, http://www.sanger.ac.uk/Projects/B_thuringiensis/).</p>
FEATURES	<p>Location/Qualifiers</p> <p>1..117923</p> <p>/organism="Bacillus thuringiensis serovar israelensis"</p> <p>/mol_type="genomic DNA"</p> <p>/db_xref="taxon:1430"</p> <p>/note="plasmid pBtoxis"</p> <p>1145..1149</p> <p>1163..2089</p> <p>/gene="pBt001"</p> <p>1163..2089</p> <p>/gene="pBt001"</p> <p>/note="Similar in part to Bacillus anthracis pXol-49 TR:Q9X319 (EMBL:AF065404) (227 aa) fasta scores: E(): 8..9e-44, 78..48% id in 158 aa"</p> <p>/codon_start=1</p> <p>/transl_table=1</p> <p>/protein_id="CAD30064.1"</p> <p>/db_xref="GI:21685411"</p> <p>/db_xref="UniProt/TREMBL:Q8KNX5"</p> <p>/translation="MSTHNLFSQGNFKVHGVKFFDGLFNITKLENERKELPSALATYIMAFPCNDIGMLPREFQIMDLAKSGIPYTTITFQVCLERKLVAEIPVGRATVEIYVALYNNHTAETDMINISLSFRIPELITOTLISLVKARDNRGIMILLETITFVSKTLENNYKHKKITIDPEITRKAPFLKEDLNRAKRYROYETLVKPKFPAVLDKDDKSSSRTTRRRQVQVITIEKFNVISSNCYEDNDKELHRPIAKCKRAVSRLKHMQALKKKKNIMTAFRQEIVIDEILYPTKKQN"</p> <p>2246..3106</p> <p>/note="IS240"</p> <p>/complement(2246..2262)</p> <p>/rpt_type=INVERTED</p> <p>2337..3044</p> <p>/gene="pBt003"</p> <p>2337..3044</p> <p>/gene="pBt003"</p> <p>/note="Similar to Bacillus thuringiensis insertion element IS240-a protein TR:Q45766 (EMBL:M23740) (235 aa) fasta scores: E(): 2..2e-92, 99.57% id in 235 aa, and to Mycobacterium fortuitum, transposase tmp tmp6100 TR:Q49185 (EMBL:X53635) (254 aa) fasta scores: E(): 1..4e-37, 48.05% id in 231 aa"</p> <p>/codon_start=1</p> <p>/transl_table=1</p> <p>/product="insertion element IS240 protein"</p> <p>/protein_id="CAD30065.1"</p> <p>/db_xref="GI:21685412"</p> <p>/db_xref="GOA:Q8KNX4"</p> <p>/db_xref="InterPro:IPR001584"</p> <p>/db_xref="UniProt/TREMBL:Q8KNX4"</p> <p>/translation="MERKNI FKKVHY QADMILITVRWYLYNLS FPDLVEMMERGLS LSTTITRWVHYGPEINERIRGLKRTYDSVDEYIRIKENMYLVAVSEGGTILDTIESKDKAKACFELKKALASPVYTKRVTYQGNKAYPAVIRELAKSIEIPYQMAPIAVKVIYNNMTIEDHRIKGRITLNLGLKSMQTAVMITGIDAMHVVKKGGLKLRQSAQNORCHIQHQLGLTA"</p> <p>2538..2933</p> <p>/gene="pBt003"</p> <p>/note="HMPfam hit to PF00665, Integrase core domain"</p> <p>3090..3106</p> <p>/rpt_type=INVERTED</p> <p>3521..4381</p> <p>/note="IS240"</p> <p>/complement(3521..3535)</p> <p>/rpt_type=INVERTED</p> <p>3612..4319</p> <p>/gene="pBt004"</p> <p>3612..4319</p> <p>/gene="pBt004"</p>
misc_feature	
repeat_unit	
repeat_unit	
repeat_unit	
gene	
CDS	

/note="Similar to Bacillus thuringiensis insertion element IS240-a protein TR:Q45766 (EMBL:M23740) (235 aa) fasta scores: E(): 3.5e-91, 99.14% id in 235 aa, and to Mycobacterium fortuitum, transposase tnpA or tnp6100 TR:Q49185 (EMBL:X53635) (254 aa) fasta scores: E(): 1.1e-37, 48.05% id in 231 aa"

/codon_start=1
/transl_table=1
/product="insertion element IS240 protein"
/protein_id="CAD30066.1"
/db_xref="GI:21685413"
/db_xref="GOA:O8KH55"
/db_xref="InterPro:IPR001584"
/db_xref="UniProt/TREMBL:O8KH55"
/translation="MEKENIPKMKHYQADMLMTWRYLYRNLSPRLVEMMERGLS LSHITIMWVHOYGBELNERIKRLKRPNDSDRVDYTIKIGENMYLYRAVDSRGNT LDFPLSKRDAKAKACPLKALASFTHTKPVYITVDGKAYPAVALREIKNEKSLPYGM PLRVKYLNNMI EODHPIKRIIRNMLGLKSMQTAVKMIAIGIEAHMVKQLKLRQ SAQONRCIADLPGLTA"
3813 . 4208
/gene="pBt004"
/note="HMPfam hit to PF00665, Integrase core domain"
4365 . 4381
/rpe_type=INVERTED
complement(4634 . 5275)
/gene="pBt005"
complement(4634 . 5275)
/gene="pBt005"
/note="Similar to N-terminus of Bacillus anthracis pXol-18 TR:Q9X2Y9 (EMBL:AF065404) (315 aa) fasta scores: E(): 4.2e-56, 84.15% id in 183 aa, and to Bacillus halodurans Bb2364 protein TR:Q9KAC5 (EMBL:AP001515) (378 aa) fasta scores: E(): 1.6e-18, 35.45% id in 189 aa, and weakly to Laccobacillus delbrueckii integrase/recombinase crt2 TR:Q48538 (EMBL:Z50864) (333 aa) fasta scores: E(): 6.3, 28.88% id in 90 aa, and to Bacillus thuringiensis resolvase tnpI SW:TNRI_BACTV (P10020) (284 aa) fasta scores: E(): 8.5, 23.88% id in 180 aa"

/codon_start=1
/transl_table=1
/product="Integrase/recombinase family protein"
/protein_id="CAD30067.1"
/db_xref="GI:21685414"
/db_xref="GOA:O8KH33"
/db_xref="InterPro:IPR004107"
/db_xref="UniProt/TREMBL:O8KH33"
/translation="MSKNINLBETLAASAVLIEKRRKSTIKRAYADIKDFYKMLRA NKLIHITSMSLEADQYTFSELEDRKYSOKTRHRIMVVLKLIHMFICIVSLDG INLSLIPOSINDNDFTIMEEKLKOTVLSKGLTERQAKYRPLIMDRANCIIVLV NYGSLQELVSLNMSHINFAINICVYKIDITLNRDILFVQLBIVAFHRSF"
complement(4985 . 5251)
/gene="pBt005"
/note="HMPfam hit to PF02899, Phage integrase, N-terminal SAM-like domain"
complement(5283 . 5286)
complement(5364 . 5507)
/gene="pBt006"
complement(5364 . 5507)
/gene="pBt006"
/note="Similar to Bacillus anthracis pXol-17 TR:Q9X2Y8 (EMBL:AF065404) (47 aa) fasta scores: E(): 2.1e-12, 68.08% id in 47 aa"

/codon_start=1
/transl_table=1
/product="putative integral membrane protein"
/protein_id="CAD30068.1"
/db_xref="GI:21685415"
/db_xref="UniProt/TREMBL:O8KH33"
/translation="MAVLKDSNYIEAMTYNEIGFPGKGI FVGLIFVVPWFSIMATFIW LCK"
complement(5370 . 5435)
/gene="pBt006"
/note="1 probable transmembrane helix predicted for pBt006

misc_feature

repeat_unit

gene

CDS

RBS
gene
CDS
by TMHMM2.0"
complement(5515 . 5519)
complement(6451 . 8160)
/gene="pBt007"
complement(6451 . 8160)
/gene="pBt007"
/note="Similar to Bacillus anthracis pXol-16 TR:Q9X2Y7 (EMBL:AF065404) (569 aa) fasta scores: E(): 0, 96.13% id in 569 aa, and to Bacillus thuringiensis pXol orf16-like protein TR:CA050562 (EMBL:AJ296638) (310 aa) fasta scores: E(): 6.5e-122, 99.67% id in 310 aa"

/codon_start=1
/transl_table=1
/product="conserved hypothetical protein"
/protein_id="CAD30069.1"
/db_xref="GI:21685416"
/db_xref="UniProt/TREMBL:O8KH31"
/translation="WGSFLIQNRSINGKTNAVVTLSKQKRVAAFEKYYSYKD TYLEDILQHMTEERADVEGFEKARSPERRAMSLAETGRKHLGYGTNADLEV FLYLHKCCFTGVIPNVTIHMWEDYKQKEFAYIOHSQYIALKLIHNLISIEH GLDGRVYIKLTHFNNEETEKANPYVYISPVIFTPDFFLSIAKKLFLDIMOOHTEF TIKSLKODERGRKTRGGMVRFLLHKYHPHOIRAVBELTALPCGNLPFKICKMO KGVNTRKRYTTLTYLISHSDPLCSKEABEQBRDPTPKATYARAKPIEAVLOENITG ELSADWKKFTHVLKYTCRQIRSVIRGLRDVDRKEGYPIKVTLLKLIHQTSOYL LDTAKKGIYPLIQHLPKERNSDREQAVFNFGALSNVSLNLI KRMFKVHALKQK FAVPVEESYVRNLYKYQETLFRKVAVDGVNLAAYIALEINREKLYGKHKERTI PSDVREWFIEAIDRLPOEKLRVIELPQPNLLEFMRTPERLVRAGVTTIAPDYLHAM EIK"
complement(8168 . 8173)
8947 . 8951
8959 . 10653
/gene="pBt009"
8959 . 10653
/gene="pBt009"
/note="Similar to Bacillus anthracis pXol-14 TR:Q9X2Y5 (EMBL:AF065404) (564 aa) fasta scores: E(): 1.6e-191,

RBS
RBS
gene
CDS

Query Match 15.6%; Score 345; DB 1; Length 127923;
Best Local Similarity 53.2%; Pred. No. 2.4e-38;
Matches 1175; Conservative 0; Mismatches 935; Indels 99; Gaps 17;

4 AATCAAAATTAAGATATACGAATATGAAATTAATGATGCAATCTCACTTATTTT 63
Db 96531 AATATGAATCCTTATCAAAATTAATAATGAATGAACATTAATAGCTTCAAAAAA 96472

64 CCGAACAAGAACAGTATGATTTCTAGTACCTTAACAAATTAATCCAAATCAACATTA 123
Db 96471 TTAATATATCTAATTAATATATACAAATTCATAGAAATATAGCCAAACATATTA 96412

124 CAAAACAAGATTAACAAGAGTGCCTCAATATGTCMAAGGAATPACAAATATGAT 183
Db 96411 CAAGTACAAATTAATANAAGATTGCTCAATATGTCMAAGGAATACAGCATATAGTGA 96352

184 AATTGAGACATTGCTAGTGCATATCAATTCGACATTAAGTGCAGTACTATTTGTA 243
Db 96351 GATTTGAAACTTTTATTTGAT-----AGTGTGAATCGAGTGCCTATACTATTTGTA 96301

244 TCCGTAATCTGTTAGCCGGATATAGTGGGCTCACTTATATCCGACCATATGAGATA 303
Db 96300 GTTGGACCGTACGACATGCTG-----TTCCGGGTTGACACACACC 96262

304 ATAGTGCTATTAATATATCTTTTGGTACCTTAATCACTGTCTTTTGGCCCGGAGAA 363
Db 96261 TTAGGACTTGCTTTAATATAGTTTTTGGTACATTAATACAGATGCTTTTTCAGCCCAAGC 96202

364 CAAGACAAAACAGTATGACACAAATTAATTAATGGAAGAAATTTTGTGATACACCG 423
Db 96201 CAATCTAACAC--ATGAGAGGACTTATATACACAAACTAAATAATATATTAATAAAGAA 96145

424 TTAACA-GAAGCATTAACAGCTTAAGCTTAACAACCTTAAAGATTTAACAATAATT 482
Db 96144 ATAGCATCAACATATATTAAGTAAATTAATTAATTAATTAACAGAGTGTTAATGTTATC 96085

QY 483 ACAAGCTATATACAGCTTATGATGGAGAAAATTAAGACTACAGCTCTCG 542
DB 96084 AGCACTTATATCACTTAAACATGGAGATATCCAAACCAACAAATATCTAG 96025
QY 543 ATACCAACCATCATGAGCTTACAAACAGCTGCCCTGACTCTTAAATATGACTTGAGAA 602
DB 96024 GATGTAAAGAA---CACAAATCCAGCTTATGATTTACATTTTCAAAATGTCATTTCCAGA 95969
QY 603 TGTTTCAATGATTTTATTCGAAATATACCTGTGTTCCACCTGAAAATTAAAGCGT 662
DB 95968 GCTTGTAAATCTGTCTCTCTTAATCTGATGATTC---GATTAATTAATCACT 95915
QY 663 ATTAATCACTTATATGCGCAAGCTGCTTATTTTCAATTTTAAATTTTATACCAAGCTGC 722
DB 95914 AGTATTTATCTATGATGCAAGAGCAAGCAAACTTATCTGATCTATTTAAATCAAGCGCT 95855
QY 723 TGAATGGCTGATGATGAGATGAGATATATCTTCACAAAATGAGAACTTAATGCTGG 782
DB 95854 CAATTTGAAAGGATTTTAAACAAATCGCAATTCGATTTATTTAGAGCTTTGCC--- 95798
QY 783 AACATCAAGATGATTTATTAACCTTTTAAAGAAATATATCTTAATATATGATTAATG 842
DB 95797 AACGCAATGATTTATTTATCCAGATTTGATTAAGCTATAGAGATTAATCACTAATTTATG 95738
QY 843 TGCAAATATCTATAGAACAGACTTAAATCTTATAGAGCAACCAATATGAAATGAG 902
DB 95737 TGTAAACAATTTATTAAGATTTAAATTTAAACGAGCGCTGATGATTAATCTTGA 95678
QY 903 TATATTTAATGATGATGAGATATATGACCATTTACTGATTTAGATTAATCACTCTCAAT 962
DB 95677 TGGAAATTAATCGAACAATA-----CAATATCGATGAAACAAATATGATCTAGCTGC 95624
QY 963 TTCTTTATATGATTAAGAAATATAGATTTCAATAGAGAAATAGAAATGAAAGCAT 1022
DB 95623 TGTATTTAGATCTGTGTCAGCTTTCTTAATTTATGATGATGATTAATTCATATAGGT 95564
QY 1023 TAAAGATGATCTCACAGAGAAATTTATACACTGAAATTAATTTGATGCTCTCTCA 1082
DB 95563 CCAATCTGAACTTACTCGAGAAATTTAT--CAGGTACTTAATCTCGAAGAAAGCCCTAT 95506
QY 1083 ACTTAGAGTCAACCCATCTAGCTAGAGTGAATATTAATTAACAGTGCAGCTTTAA 1142
DB 95505 AAATATATGATCTTCAATATCAAGAGATTCATCTAGACGTAGAC--GCATTTATTTA 95448
QY 1143 ATTAATTTTATTTAGAACAAATTTATTTTATATACAGAAATACAAATTTGCGAAATCG 1202
DB 95447 CTGGCTGATTTCTTTGAA-----TTTTATGAAAAAGCGAAACTACTCTAATTAATTT 95393
QY 1203 TTTAGTTGATTTCTAATCGTATGACCTACTTATAGCAATCTATATCTGAAATCTTT 1262
DB 95392 TTTTACCGCCATTAATATATGTTTATATACACTGATATATATATCCAAAATCTAG 95333
QY 1263 ATATGAGAAAGAAAGAGTTTACCCACAACAAAACATAAGACATTTGAAATCTTTAA 1322
DB 95332 TGTTTTGGAAATCACATGTAATGATTAATTAATCTTGGTTTGGCAAAATAT 95273
QY 1323 AGTTTCAATTTGTAATGATAGCAATCACTCTGTTTCCCTATATCAACCACTTTAT 1382
DB 95272 TTATATTTTATTAATATGTCATAGCTTATGATTAATTAATATGATTAATTAATTA 95213
QY 1383 AATTATCAATTTGAACTTTATTTAATGCTCATCTTACACACACTCAAAATTTTCAAGC 1442
DB 95212 TATTTAGTAAATGATTTTATTAATGATGTAAGTACTTTTGGAGAAAGAACTTAC 95153
QY 1443 AGAGAGGCT-----TTATCAATTTATCAAAACCAACTTTTTCATTTCTAGAAA 1496
DB 95152 AGCAGGATCTGGCAATTAATCTTATGATTAATTAATTAATTTTTCGGTTATCAAAATCT 95093
QY 1497 AAAAGATGCAATCTAGTTATTTGATCCAGGTGTTTCAACCAATTTATTAATTAATGCA 1556
DB 95092 TAAAGAAAGAGATCAAGAAACCCATACCTTTTCCAAATATGATTAATTAATGCA 95033
QY 1557 TATTTATCCATTTTTCATTTATTTACTTATTTCTATGATGATGATTAATTAATTAAT 1616

DB 95032 TATTTATC-----ATTATTAAGTCTTGTATCTCTGCAACATAT 94990
QY 1617 ATTAGATACAGGTGTA--TTAGATGACACACAGTATGTTATATATATATGCAATA 1674
DB 94989 AAAATCAATGATATATGTTTCTTGGACACACTTATGTTTATCTTAAATATACATTT 94930
QY 1675 TCAGATTAATTAATTAATCAATGATCCGCAATCAAGATTAATCAATCTGATACAACTCT 1734
DB 94929 TATACACTTTAATCTACCAATTTCCAGCTGTGTAAGACGAATTCATCTTGGAGCTGCTCT 94870
QY 1735 AAGTATTTAAGACCTGCTGATACAGAGAACTTGGTTTATTAACAAAGTCAAGG 1794
DB 94869 AAGGTGTTCAAGACCTGCTGATACAGAGGAGATTTAATGATTTCAAAAGATCA---- 94814
QY 1795 CGTTTGAATTAATCATGTAAGTGAATCTCTAATCTTACACATCTTATTTATTTAGACTTCA 1854
DB 94813 --TTTCAAAATTAATCTTCAACTCAACTCAAAATTTTCAACATCTGATTTATTAAGATTCGT 94756
QY 1855 TATGCTAACAAATGATGCTGCAAAATCTCTTCTAATATATCTTACAAATACAGAGATA 1914
DB 94755 TATGCTTCAATGGAAGCGCAAAATCTCGAGCTGTATTAATCTTATGATATCCAGAGGG-- 94698
QY 1915 ATAGGAATACCACTCTCAACGATCAACACACTTTTCTGATCAAAATTAATTAATTTA 1974
DB 94697 -TAGCAAGACTGGATATGACACTCAACCCACTTTTCTGATCAAGATTAATGCAATTTA 94639
QY 1975 CAATACGAGATTTTGGGTATTTTCAATTTTCAAGTACAGTAACTTATCTTTAAATCA 2034
DB 94638 AAATTAATTAATTTTCTGATCTTATGAAATTTTCTTACAGAGTGAATTTGCTCCAAATCA 94579
QY 2035 AACATACATTTATTTATTTAATCTGTCAGATGAT--CAATTCATTTATATCATTTGAT 2091
DB 94578 AACATATCTCTGTGTTTATCTGTGCGATGATATATCAAAACACAGATATCTTATGAT 94519
QY 2092 AAAATGAATTTTATACCAATTTACTCTCTATATGACCAAAATAGAGAAACAAATTA 2151
DB 94518 AAAATGAATTTTCTGCAATTTACTCTGTTATAGAGAGATGAGAGAAACAAATTA 94459
QY 2152 GAAATCTCCAAACAAAAATTAATACATTTTTCACAAATCATACAAATA 2200
DB 94458 GAAACGATACACAAATTAATTAATACATTTTATGCAAAATCTTATTAATA 94410

RESULT 10
BTH251977
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

REFERENCE
1
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES
source
1.6009
/organism="Bacillus thuringiensis serovar medellin"
/mol_type="genomic DNA"
/isolate="163-131"
/sub_species="medellin"

BTH251977 6009 bp DNA linear BCT 22-DEC-2001
Bacillus thuringiensis subsp. medellin cry29a gene for Cry29a protein.
AJ251977.1 GI:17977978
cry29a gene; Cry29a protein.
Bacillus thuringiensis serovar medellin
Bacillus thuringiensis serovar medellin
Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus cereus Group.

Delecluse, A. and Orduz, S.
Characterization of two new mosquitoicidal toxins, Cry29a and Cry30A, from Bacillus thuringiensis medellin
Unpublished
2 (bases 1 to 6009)
Delecluse, A.
Direct Submission
Submitted (22-DEC-1999) Delecluse A., Bacteries & Champignons Entomopathogenes, Institut Pasteur, 25, rue du Dr Roux, 75724 Paris Cedex 15, FRANCE
Location/Qualifiers

/db_xref="taxon:79672"
 /country="Colombia"
 /gene="cry29Aa"
 /gene="cry29Aa"
 /gene="cry29Aa"
 /function="putative mosquitoicidal toxin"
 /codon_start=1
 /evidence=experimental
 /transl_table=11
 /product="Cry29Aa protein"
 /protein_id="CAC80985.1"
 /db_xref="GI:19797979"
 /db_xref="GOA:O8VNX2"
 /db_xref="UniProt/TREMBL:O8VNX2"

translation="MNSFQNKVEILINAPSSITSPNNYSKYPIANTLNMONTNY
 KDMTMCALDNNLKSINPEANLQNSLVGLPITTAIISLSAPITGGTSLAAGTAIA
 AAIIPIWPSQENNLDPDKLAISEATLYSPQQRREDAITLSELSKDSVXFENAF
 FWINNNSNTITTVREPOEVNRFVGSVAFPRANRBPILSTYAQAARLHLARD
 GITYAEKNLSRQGDMDGDLLEYFNKCYEIEHCLIKMYNESLSLKSVCANLEY
 NOYRPLASVLDVLSFSSYDPLRYKERLSVEILTRKLYDNPINMYHRSLSLEADSK
 YLREPLFQIATLTFPSNIIPYVMGHNTYRYSIPDKIPARSRGSSYIDKYPVI
 PNDKSIYKIRAYDNDHNGLEFNWTFGWDEGKDOIQKIIQSSSTIYIKNCTHRLADV
 ISHLDENKRCYSAFAMSTTISLENEIKNDIITQIPAKAYQAGVQSVIKGPHGTG
 DLINKSDYLRISSQHLNVTYKRYFIRIYATNGSLNTRPIINITIGMTPGQWVLD
 NTFSGYSLNEYQNGYKFEFLKEVTLNPNOSISLTNRSQNSIILLDRIFLPI
 TPSIRKSEQQLBKNGKTVNLFN"

ORIGIN

Query Match 10.1%; Score 222.6; DB 1; Length 6009;
 Best Local Similarity 65.4%; Pred. No. 3.2e-21;
 Matches 375; Conservative 0; Mismatches 189; Indels 9; Gaps 3;

OY 1633 TTAAGATGACACAGAGTGTGATAGATTAATGCAATTCAGTAAATATTACA 1692
 DB 2031 TTTGATGACACATCTACTACAAATTTCTTGAAATGAAATGAAATGATTTTCACA 2090
 OY 1693 ATGATCCAGCAATCAAGGTAAATCTTGATACAACTCTAAGTAAATGAGGACCT 1752
 DB 2091 CAAATCCCTGCTGTAAAGCCATCTAGTGTGCAATCAAGTCAATTAAGTCCA 2150
 OY 1753 GGTCTACAGAGGAACTGGTTATTATTCAAAGTCAAGGGCGTTAGAAATTAACATGT 1812
 DB 2151 GGTCTACGTGGTGAAGCTTAATTAATTAATCAAAATGATTAATAGTATATCATGT 2210
 OY 1813 GAA---ACTCCATATTCACAACTTATTTCAATGACTCGATAGCTCAATAGT 1869
 DB 2211 CAACACTATCTAATGTTACAAAAAATATTTGTAGAAATTCGCTACGCTACAAATGGT 2270
 OY 1870 GCTGGAATATCTCTCTAATATATCTCTTACAAATGAGAGTATGGAATACACCT 1929
 DB 2271 TCTCTAATATCTAGGCCCATTTAATTAATTTACAAATTCAGG---AATGACACACAGGT 2327
 OY 1930 CAACGACTCAACACACTTTTCTGTGCAAAATTTAATTAATTTCAATACGAGATTTT 1989
 DB 2328 ATGTAATCTGCAATTAATCTTTTCTGTGACAGGTATAGCAATTTAGAAATTAAGAAATTTT 2387
 OY 1990 GGGTATTTTCCAAATTTCCAAAGTACATTAATCTTTAATGCAACATACATTTATA 2049
 DB 2388 GGGCTTAAAGATTTCTAAAGAGGTAAATTAATTCCAAAATCAAGATATATCTTTACA 2447
 OY 2050 TTTAATCGGAGATGTA---TCAAATTCAAATTTTAATCATGTGATTAATGAAATTTATA 2106
 DB 2448 TTTAATCGTTGTCATCAAAATTCAAATTTTAATCTCTGATGATGATTAATGAAATTTCTA 2507
 OY 2107 CCAATTAATCTCTCTATGACACCAAAATAGAGAAAAAATTAAGAAATCTATCCAAACA 2166
 DB 2508 CCAATTAATCTCTCTATGACAGAAAGTAAAGAACAAATTTTAGAAAAAATCAGAAA 2567
 OY 2167 AAAATTAATCACTTTTTCAGCAATCATACAAA 2199
 DB 2568 ACGGTAAACAAGTATATTTTCAATTAATAATAA 2600

RESULT 11
 BTPBTOXIS
 LOCUS
 DEFINITION
 ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 PUBMED
 REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 COMMENT

BTPBTOXIS 127923 bp DNA linear BCT 16-APR-2005
 Bacillus thuringiensis subsp. israelensis plasmid pBtoxis.
 AL731825
 AL731825.1 GI:21685410
 Bacillus thuringiensis serovar israelensis
 Bacillus thuringiensis serovar israelensis
 Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus; Bacillus
 cereus group.
 1 (bases 1 to 127923)
 Berry, C., O'Neill, S., Ben-Dov, E., Jones, A.F., Murphy, L., Quail, M.A.,
 Holden, M.T., Harris, D., Zaritsky, A. and Parkhill, J.
 Complete sequence and organization of pBtoxis, the toxin-coding
 plasmid of Bacillus thuringiensis subsp. israelensis
 Appl. Environ. Microbiol. 68 (10), 5082-5095 (2002)
 2 (bases 1 to 127923)
 Parkhill, J.
 Direct Submission
 Submitted (19-APR-2002) Submitted on behalf of the pBtoxis
 sequencing team, Sanger Centre, Wellcome Trust Genome Campus,
 Hinxton, Cambridge CB10 1SA E-mail: parkhill@sanger.ac.uk
 Notes:
 Details of pBtoxis sequencing at the Sanger Centre are available on
 the World Wide Web.
 (URL, http://www.sanger.ac.uk/Projects/B_thuringiensis/).

FEATURES

source
 1..127923
 /organism="Bacillus thuringiensis serovar israelensis"
 /mol_type="genomic DNA"
 /db_xref="taxon:1430"
 /note="Plasmid pBtoxis"
 1145..1149
 1163..2089
 /gene="pBt001"
 1163..2089
 /gene="pBt001"
 /note="Similar in part to Bacillus anthracis pXol-49
 TR:Q9X319 (EMBL:AF065404) (227 aa) fasta scores: E():
 8.9e-44, 78.48% id in 158 aa"
 /codon_start=1
 /transl_table=11
 /protein_id="CAD30064.1"
 /db_xref="GI:21685411"
 /db_xref="UniProt/TREMBL:O8KNX5"
 /translation="MSTHNFQONFKVGVFDFGLFNIKLENERKELPSSALATY
 IMLPECDIDGLPREQIMDLAKSGIPYTTITGPGVCEKRLVREIPIGNRTVYE
 IVDYALXNHTAETMDMINISLSEYRIPVLTITITLSLVKARDNRMILLETINT
 FSKRTGHEHTKIKEDFEITRKMAFLKELKRNKAKROYEIRYPIFNADVLDKDK
 SSESRIYRVRQVQVQVIEKFNVPSSCVLENDEKELHPIACRKAASVRLKMGQ
 ALRKXENIMTAFROEIVDIETVLPKOKN"
 2246..3106
 /note="IS240"
 complement(2246..2262)
 /rnc_type="INVERTED"
 2337..3044
 /gene="pBt003"
 2337..3044
 /gene="pBt003"
 /note="Similar to Bacillus thuringiensis insertion element
 IS240-a protein TR:Q45766 (EMBL:M3740) (235 aa) fasta
 scores: E(): 2.2e-92, 99.57% id in 235 aa, and to
 Mycobacterium fortuitum, transposase tnp tnp6100 TR:Q49185
 (EMBL:X53635) (254 aa) fasta scores: E(): 1.4e-37, 48.05%
 id in 231 aa"
 /codon_start=1
 /transl_table=11
 /product="insertion element IS240 protein"
 /protein_id="CAD30065.1"
 /db_xref="GI:21685412"

repeat_unit
 repeat_unit
 gene
 CDS


```
/db_xref="GOA:Q8KXK4"
/db_xref="UniProt:IPR001584"
/translation="MEKENIFKMGHYQADMLTMVRYLYRNLSPFDLVEMMERGLS
LSHTTIRMVQVGPBELNERTIRKILKRTNDSMRVDEYIKIKGNMYLYAIVDEGNT
LDPLYSKKRDAAKCFPKKALASFHTKPVITVDNKAYPVARELKNKKSIPYQM
PLRVKYLNNNIEDHPIKRIILMGLKSMQTVAMGIAGIEAMHVKQGLKLRQ
SAQNQRNCIHOLFGLTA"
2538..2933
/misc_feature
/gene="pBc003"
/feature="HMPFam hit to PF00665, Integrase core domain"
3090..3106
/repeat_unit
/rpt_type="INVERTED"
3521..4381
/feature="IS240"
/complement(3521..3535)
/repeat_unit
/rpt_type="INVERTED"
3612..4319
/gene="pBc004"
3612..4319
/gene="pBc004"
/feature="Similar to Bacillus thuringiensis insertion element
IS240-a protein TR:Q45766 (EMBL:M3740) (235 aa) fasta
scores: E(): 3.5e-91, 99.14% id in 235 aa, and to
Mycobacterium fortuitum, transposase tmpr or tmpr100
TR:Q49185 (EMBL:X53635) (254 aa) fasta scores: E():
1.1e-37, 48.05% id in 231 aa"
/codon_start=1
/translation="conserved hypothetical protein"
/protein_id="CAD30066.1"
/db_xref="GI:21685413"
/db_xref="GOA:Q8KH55"
/db_xref="UniProt:IPR001584"
/translation="MEKENIFKMGHYQADMLTMVRYLYRNLSPFDLVEMMERGLS
LSHTTIRMVQVGPBELNERTIRKILKRTNDSMRVDEYIKIKGNMYLYAIVDEGNT
LDPLYSKKRDAAKCFPKKALASFHTKPVITVDNKAYPVARELKNKKSIPYQM
PLRVKYLNNNIEDHPIKRIILMGLKSMQTVAMGIAGIEAMHVKQGLKLRQ
SAQNQRNCIHOLFGLTA"
3813..4208
/misc_feature
/gene="pBc004"
/feature="HMPFam hit to PF00665, Integrase core domain"
4365..4381
/repeat_unit
/rpt_type="INVERTED"
/complement(4634..5275)
/gene="pBc005"
/complement(4634..5275)
/gene="pBc005"
/feature="Similar to N-terminus of Bacillus anthracis pxc1-18
TR:Q9X2Y9 (EMBL:AF065404) (315 aa) fasta scores: E():
4.2e-56, 84.15% id in 183 aa, and to Bacillus halodurans
Bh2364 protein TR:Q9KAC5 (EMBL:AP001515) (378 aa) fasta
scores: E(): 1.6e-18, 35.45% id in 189 aa, and weakly to
Lactobacillus delbrueckii integrase/recombinase oriz2
TR:Q48538 (EMBL:Z50864) (333 aa) fasta scores: E(): 6.3,
28.88% id in 90 aa, and to Bacillus thuringiensis
recombinase tnp1 SW:TNRI BACTU (P10020) (284 aa) fasta
scores: E(): 8.5, 23.88% id in 180 aa"
/codon_start=1
/translation="1"
/product="integrase/recombinase family protein"
/protein_id="CAD30067.1"
/db_xref="GI:21685414"
/db_xref="GOA:Q8KXK3"
/db_xref="UniProt:IPR004107"
/db_xref="UniProt:TEMBL:Q8KXK3"
/translation="MSKINLEBTLAASALILKGRKOSTIKYAVDIDPFYKMLA
NKLIHLKMSWELSBADYQYFSELEBRKXSQKTRIRIIVWLKMLPFCISPLDG
INLSLIPDSINDNDFTIMEBKLLKQVLSVTKGLTIRIOWALKMLPFCISPLDG
NYGLSGLVLSINSHINFAINICVYKKIDTLNRDILFVQLEIVAFHRTSF"
/complement(4985..5251)
/misc_feature
/gene="pBc005"
```

```
/note="HMPFam hit to PF02899, phage integrase, N-terminal
SAM-like domain"
5286)
complement(5283..5286)
/gene="pBc006"
complement(5364..5507)
/gene="pBc006"
complement(5364..5507)
/gene="pBc006"
/feature="Similar to Bacillus anthracis pxc1-17 TR:Q9X2Y8
(EMBL:AF065404) (47 aa) fasta scores: E(): 2.1e-12, 68.08%
id in 47 aa"
/codon_start=1
/translation="1"
/product="putative integral membrane protein"
/protein_id="CAD30068.1"
/db_xref="GI:21685415"
/db_xref="UniProt:TEMBL:Q8KXK2"
/translation="MAVLKDSNYIEMNTYIEISFPGKIFWGLFVDFMSIMATFIW
LCK"
complement(5370..5435)
/gene="pBc006"
/feature="1 probable transmembrane helix predicted for pBc006
by TMHMM2.0"
complement(5515..5519)
/gene="pBc007"
complement(6451..8160)
/gene="pBc007"
/feature="Similar to Bacillus anthracis pxc1-16 TR:Q9X2Y7
(EMBL:AF065404) (569 aa) fasta scores: E(): 0, 96.13% id
in 569 aa, and to Bacillus thuringiensis pxc1 orf16-like
protein TR:CA650562 (EMBL:AJ296638) (310 aa) fasta scores:
E(): 6.5e-122, 99.67% id in 310 aa"
/codon_start=1
/translation="1"
/product="conserved hypothetical protein"
/protein_id="CAD30069.1"
/db_xref="GI:21685416"
/db_xref="UniProt:TEMBL:Q8KXK1"
/translation="MGSLFIQNRSLNGTAAVAVVTLKSWQFVAAFEKYYSYKD
TYLEDLIQHTYERIALADVGEFVEKAPSPSRAMSTLAKTGKRLKGLYITNADLEV
FLYLHKCFPTNGVLPVNTLIHMMEDYKQYKEPFIQHSQFIYALKLNLIHISIN
GLGRYTLIKLTHPMNSETKANPYVYISPIFKDPKLSIAKKFLDIAMOOHTCT
TLRSLDKODRGNKTHFGMYRFELKKYHQIRAYBELTLLPCNGDLFICTCKQ
KGVKHTKTYTLTSLHSDPLSKAGSEBHRPFPKATYARAKFTAVLDBMNTG
ELSDANKFTHLVKTYCHQRQIRSVIRGLRDMVRKGVPRKYVTLAKLHQTSQOI
LDTAAKGGIYPLIAOHIPKERNDSRQAVNFGIHSMSYLNHIKMPKRVNHLXOK
FAVPTBESYRYVLYKQSEETLPKRYAVDGVNLYHAYIAIEIMREKLKRGKERTI
PSDVREMFETAIIDKLPQEKRVIELPQGFILSPMTFERLVRAAGVITAPDQVLHAM
EIK"
complement(8168..8173)
RBS
8947..8951
RBS
8959..10653
/gene="pBc009"
8959..10653
/gene="pBc009"
/feature="Similar to Bacillus anthracis pxc1-14 TR:Q9X2Y5
(EMBL:AF065404) (564 aa) fasta scores: E(): 1.6e-191,
Query Match 8.64; Score 189.6; DB 1; Length 127923;
Best local Similarity 50.04; Pred. No. 2,7e-17;
Matches 1107; Conservative 0; Mismatches 924; Indels 183; Gaps 18;
```


OY	181	GATAATTCCAGACATTTGGCTAGGCTGATACAAATTTGCGCACTAGTGTGACAGTACTATT	240
Db	39235	AATTAATGCGGGAAATTTTGCTAGTTCTGAAACTAATGTGTGAATAGTGTGAGGATTAATT	39294
OY	241	GTATCCGACTCTGTATGCGCGATATAGTGGGCTCACTTCTATATCCGAGCGATAGGA	300
Db	39295	GTACTAGGAACCTAATGTTAGAGCTTTTGGCGCCCT-----	39330
OY	301	ATATATAGTGTATATATATATCTTTTGGTACCTTAATCATCTGTCTTTGGCCCGGGGA	360
Db	39331	GCTTACTGTGAGGTAAATATCTTTTGGGACTTTGTGGCAATCTTTTGGCAAGATCT	39390
OY	361	GAACAAGACAAAACAGATGAGCACAATTAATTAATAATGGAGAAATTTTGTGTATACA	420
Db	39391	GAACCTG---CAATGTTTGGCAGAAATTTGTTAACATCGAGAGAAAGCCCTATCAAGA	39447
OY	421	CCGTTAAACAGAAAGCATTAACAAGCTAAAGTTACAACTTTAGAGAAATTTAGACAAATA	480
Db	39448	ATAGATAAAAACATTAATTAATGTACTAACTT-----CTATCGTAAACCTATTAATAAAT	39501
OY	481	TTACAAGGCTATATAACAGACTTAGATAGATTGGAGAAATTTAAAGACTACAGACTCT	540
Db	39502	CACTTGTATATATATCAAGAAATTTTGCATTAATTTGGAGCGACACAGTACACGCTAAT	39561
OY	541	GGATTACACACCATCATGCACTTAACAACAAGCTGCTTGACTCTTAATAATACGATTTGAG	600
Db	39562	GCTAAACAGTA-CATGATCTCTTAACCTTAAGAACCTATATATGATTAAGATTTACA	39620
OY	601	AATGTTCAACAATGTTTATTTCCAGAAATACCTGTGTTCCAACTTGAACTTATTAACG	660
Db	39621	TATGTTAAAA-----ATATCTGTACATACGAATA	39651
OY	661	CTATTACTACCTATTTATGCGCAAGCTGCTAAATTTTCATTTAAATTTATTAACAACAGGT	720
Db	39652	CCAAACATCCTCGATATGCACAAAATAGCTACTTGGCACTTGAAATTTATTAACAATGCT	39711
OY	721	GCTGAATTTGGCTGATGAAATGCAATATATCATCTTCACAAAATTTGAACCTTAATGCT	780
Db	39712	GCTACCTATTAACAATATATGCTGCAAAATCAAGATAT-----AATCCAAAGTACTTTC	39765
OY	781	GGAACTACAGATGCTATTTAATACCTTTTAAAGAAATATATACCTAAATATATAGTATCT	840
Db	39766	AATTCATCTAATTAATCTATTCAGGGCTATTTTAAACGTTAAATACAGAAATATCTGCTAT	39825
OY	841	TGTGCAATATCTATATAGAACAGCACTTAATAAATCTTAGAGACGAAACCAATATGAAATG	900
Db	39826	TGTTATCAAAAGTACATGAGAGCTTACATATATATATATATATATATATATATATATATAT	39885
OY	901	AGTATATTTAATGACTATGCAAGATATATGACCACTTATCTGATTTAGATATCAATCTCTCA	960
Db	39886	AATATGTATATATCTTACCGTTTGAAGATCACTTAACGTGTATATATCTTATTTGCTAT	39945
OY	961	TTTTCTTATATGATTAATAAAGTATATAGATTCAATATGAGAGAAATAGAAATTAAGGC	1020
Db	39946	TTTCCAAATTTATGACCAAGAAAATATC-----AATAGCA	39981
OY	1021	ATTAAAGATGAACTCAACAAGAAATTTATCAACTGAATAAATTTGATCGTCTTCT	1080
Db	39982	GTTAAATCTGAACCTTATCAAGAAAGTTTATACAAAGTTTA-----	40022
OY	1081	CAACTTGAAGTCAACCACTTATGCTACGATGGAATATATTTTACAGTGCAGTTT	1140
Db	40023	-----TTCAGATCATTTTGAACCACTTACAGAACTTGAATAATGATTAATCTGAATCTCT	40077
OY	1141	AAATTAATTTTCATTTTGAACAATTTATTTTATATACAGAAATTAACAATTTCCGGAT	1200
Db	40078	ACATTAATTTACTTGATTAACCAAGGCGTTTTTATACAGAAATTC--TGAAGCAT	40134
OY	1201	CGTTAGTGGTATTTCTAATCGATGACACTTAATATATATATATATATATATATAT	1260
Db	40135	CTTGAATCTTAATGATATTTTCTTTTACGGTAAACAGATGCGCTTTTACATATCTAAT	40194

QY	1261	TTATATGAGAAAAGAACGGTTCACCAACAABAAACATTAAGACATTGGATCTTAT	1320
Db	40195	GATGATCGCACTAATATCTGGGAGCGGCTTCATGGAATATATTTTTCAGACACATCC	4025
QY	1321	AAAGTTTCAATTTGATCTGATAGACAATTCACCTTCGTTCCCTATTCACGACACTT	1380
Db	40255	AAAGTATTTCTTTTATAGAAAACAACCTATGATAGGTCCGAATTTG--CAGCATTA	40311
QY	1381	ATATTAATCAAAATGAACTTTATTTAAATGGCTCATCTAACACACACTCAAAATTTCA	1440
Db	40313	GAGAGTACTGAGATATATATATGAAATGATATTTTTTGGAATAGCAGGAAGTATTC	40377
QY	1441	GCAGAGGGCTTATCTAATTTATCAAAACAACAACCTTTTTCATTTCTAGAAAAAA	1500
Db	40373	G-----ATATTCATCCAAATTCACAAATGAAAAATATATATAAGAACT-----	40411
QY	1501	GACTGCAATCTAGTATTTGATCCAGGTTGTTCCAAACTTTATACTATAGCATATT	1560
Db	40417	-----GATCTTATATGATTTCCAAACCAACATGAAAAATATAAGATATGGTCATCT	40477
QY	1561	TTATCCATTTTTCATTATTTACTTATTCCTATGTGATTTGAACTTACGCTACAAATTTA	1620
Db	40471	CTATCGATATATAAAACGTAAATTTAT-----ATATTTTCAGTAGTTAGAGAAGA	40522
QY	1621	GATACAGGTGATTAGATGAGACACAGTAGTGTGATGATATATATGAAATATCATGAT	1680
Db	40522	AGAAAGTGTGATTTAGTTAGTGGACACATAGTGTGATTTCCAAATTAACAATAGATTTA	4058
QY	1681	AAATTAATTTACATATGATCCGACGATCAAAAGTAAACAATCTTGATACAACTTAAGTA	1740
Db	40582	GATATACATCACCAAAATCCAGCGCTTAAGCTTTGAAAGTAACTTGATTTGAAAAATT	4064
QY	1741	ATTGAGAGACCTGTCATACAGAGAGAACTTGTTATTTACAAAGTCAAGGCGTTTA	1800
Db	40642	GTGAAGAGTCTGTGTCACAGGGGAGACTTGGTATTTCTTAAGATAGTAGAATTTT	4070
QY	1801	GAAATTCATGTGAAACTCTCTAATTCACACATCTTATTTCAATAGACTTGATATGCT	1860
Db	40702	AGAGTTGATTTTTTAAAAAAGTTTTCGCAAAATACAACTAGCTGA--TTCTTTATGCT	4075
QY	1861	ACAATAGTGTGGAATTAATCTCTCTTAATATATATCTTTCATATACGAGGAAATAGTA	1920
Db	40759	ACTATATCTCCAAAGACACACAGT--ATTCTTAACCGAATATGATATCTATAGTGTGAGC	40811
QY	1921	ATACCACTCAACGACTCAACAACACTTTTTCTGGTACAAATTTATATATATTTACAATAC	1980
Db	40817	TCCTATATCAACTTCCGCAAAAC-----CAATGCTACAGATTTAATCATAT	40866
QY	1981	GGAGATTTGGGATTTTCCAAATTTCCAAATGACGT-----AACATTACCTTTAAT	2031
Db	40867	GCAATATTTGGATATGTACATTTCCAAAGACAGTTCCAAATTAACATTTGAAGAGTA	40922
QY	2032	CGAAACATACATTTATATTTAATCGGAGAGATGATCAAAATCAATTTAATCATTTGAT	2091
Db	40927	GACCTTTATTTATGACCTTATATGTGTACCAATCATTCATATATATATATTTGAC	40988
QY	2092	AAAATTAATTTATACCAATTAATCTCTCTATGCAACCAAAATAGAGAAAAACAATA	2151
Db	40987	AAATTCGAATTTATTTCAATTCACATCTGTATTTAGATTTATACAGAGAAATATATA	41044
QY	2152	GAATCTATCCAAACAAAATTAATATCATTTTTACAAATCATATCAAAAAACATT	2205
Db	41047	GAATAAAACACAGAAATATGGAATGATTTATTTGTATATTTAAACAAAGTTCTT	41100

RESULT 12					
LOCUS	BACMSOB	3753 bp	DNA	linear	BCT 26-APR-1993
DEFINITION	B. thuringiensis insecticidal endotoxin gene, complete cd.				
ACCESSION	M12662				
VERSION	M12662.1				
KEYWORDS	GI:143228				
SOURCE	Bacillus thuringiensis				

Db	2351	CTATCGTATATAAAACTGATTAATTTAT-----ATATTTTCAGTGTGTAGGAAAGA	2401
Oy	1621	GATACAGGTGTAATTAGATGACACACAGTAGTGTGATATATATATATGCAATATTCAGAT	1680
Db	2402	AGAAAGATGTCATTAGTTAGTTGGACACATCTAGTGTGTGATTTCCAAAATATACATATGATTTA	2461
Oy	1681	AAATATATTAACATATGATCTCCAGCAATCAAAAGTATCAATCTTGATTCAAACTCTTAAGTA	1740
Db	2462	GATACATATCACCCAAATTCACGCTCTTAAACCTTTGAAGGTGAAGTTCGATTTGAAAAAT	2521
Oy	1741	ATTGAAGGACCTGTGTCATACAGAGAGAAACTTGTGTTATTTATACAAAGTCAAGGCGTTTA	1800
Db	2522	GTGAAGGTCTGTGTGTCACACAGGTGAGACTTGTGTAATCTTAAAGATAGTATGATTTT	2581
Oy	1801	GAAATTCATGTGAAACTCTTAATTTCTACACATCTTAATTTTCATTAGACTTCGATATGCT	1860
Db	2582	AGAGTTAGATTTTTTAAAAAATGTTTCTCGACAAATATCAAGTACGTA---TTGTTATGCT	2638
Oy	1861	ACAAATGTGTCTGGAAATATCTTCTTAATATATCTTTACAAATACAGAGTAAATAGA	1920
Db	2639	ACTATATGCTCCAAAGACAAACAGT--ATTCTTAACCGAATAGATATTAATGTGTGAGC	2696
Oy	1921	ATACCACTCAACGACTCAACACACTTTTCTGTGTAACAATTAATTAATTAACAATAC	1980
Db	2697	TCCCTAGTACACTTCCGCGCAAAAC-----CAAAATGCTACGATTTAACATAT	2746
Oy	1981	GGAGATTTTGGGTATTTTCCAATTTCCAGTACAGT-----AACATTAACCTTTAAAT	2031
Db	2747	GCAATTTTGGATATGTATGATCAATTTCCAAAGAACAGTCCAAATTAACATTTGAAGAGAA	2806
Oy	2032	CGAAACATCACTTAATTAATTAATATGTCGACAGATGTATCAAAATCAATTTAATCATTTGAT	2091
Db	2807	GACACTTATTAATGAACCTTAATATGTGTACCAAAATCATCATTAATATATATATATGAC	2866
Oy	2092	AAATTTGAATTTTATACCAATTAATCTTCCCTATGACCCAAATATAGAGAAAACAATAA	2151
Db	2867	AAATTCGAATTTATTCCAATCACTCAATCTGTATTAATTAATTAACGAAACAAAATATA	2926
Oy	2152	GAACTATCCCAACAAATAAATAATACATTTTTCACAAATCATACAAAACACTT	2205
Db	2927	GAAAAACACAGAAATATGTGAATGATTTATTTGTATTAATTAACAAAGTTCTT	2980
RESULT 13			
LOCUS	DO078744	434 bp DNA linear BCT 11-JUN-2005	
DEFINITION	Bacillus thuringiensis strain LDC-9 cry4A insecticidal protein (cry4A) gene, partial cds.		
ACCESSION	DO078744		
VERSION	DO078744.1	GI:68348788	
KEYWORDS			
SOURCE			
ORGANISM		Bacillus thuringiensis Bacillus thuringiensis Bacteria; Filimicutes; Bacillales; Bacillaceae; Bacillus; Bacillus cereus group. 1 (bases 1 to 434) 2 (bases 1 to 434) Mahalakshmi A., Sujatha, K. and Shenbagarathai, R. PCR analysis and molecular characterization of cry4A gene of indigenous Bacillus thuringiensis LDC-9 unpublished	
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
FEATURES			
source			

gene	CDs	ORIGIN
<1..>434 /gene="cry4A"	<1..>434 /gene="cry4A" /note="cry4A delta endotoxin; contains putative conserved domain of delta endotoxin"	
Query Match Best Local Similarity 70.7%; Pred. No.1.5e-15; Matches 275; Conservative 0; Mismatches 108; Indels 6; Gaps 2;	8.3%; Score 184.2; DB 1; Length 434; 0; Mismatches 108; Indels 6; Gaps 2;	
Dp	1815 AACTCTTAATTTCAACAAATCTTATTTTCATTAGACTTCGATGTCACAAATGGTCTCG 3 ATCATCAATTTTCACAAATCGATATTTTATAGAAATTCGTATGCTTCAATATGGAAGCGC 62	1874 1875 AAAATCTCTTCCTAATATATCTCTTACAAATACGAGAGTAAATGGAATACCACTCAACG 1934 63 AAAATATCTCGAGCTGTATTAATCTTACTATATCCAGAGGG--TAGCAGAACTGGGTATGCG 119
Dp	1935 ACTCAACAAACACTTTTCTGTGTCACAAATTAATTAATTTTCAATATACGAGATTTTGGGTA 1994 120 ACTCAACCCCACTTTTCTGTGTCACAAATTAATTAATTTTCAATATACGAGATTTTGGGTA 179	1995 TTTCACAAATTTCCAAAGTACAGTAACTTATACCTTTAAATCGAAACATACATTTATTTAA 2054 180 CTGAGATTTTCTTAACAGAGGTAATTTGCTCCAAATCAAAACATATCTGTGTGTTAA 239
Dp	2055 TCGTCGAGATGTAT--CAAATTCATTTTAATTCATGATTAATTAATTTTCAATATTTTCAAT 2111 240 TCGTTCGAGATGTATTAATTCACAAACAGTAACTTATTTGATTAATTTTCAATATTTTCAAT 299	2112 TACTTCTCTATGACCAAAATATGAGAAAACAAAAATTTAGAACTATCCAAACAAATAT 2171 300 TACTCGTTCATTAAGAGATGATGAGAAACAAAAATTTGAAACAGTACACAAATATAT 359
Dp	2172 AAATACATTTTTCACAAATCATACAAA 2200 360 TAATACATTTTATGCAATCTTATTAATA 388	
RESULT 14		
E00614	3756 bp DNA linear PAT 29-SEP-1997	
LOCUS	E00614	
DEFINITION	DNA encoding a polypeptide having insecticidal activity(BTI endotoxin).	
ACCESSION	E00614	
VERSION	E00614.1 GI:2168893	
KEYWORDS	JP 1986005098-A/1.	
SOURCE	Bacillus thuringiensis	
ORGANISM	Bacillus thuringiensis Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus; Bacillus cereus group. 1 (bases 1 to 3756) Aran,E.U. and Toomasu,J.P. INSECTICIDAL POLYPEPTIDE Patent: JP 1986005098-A 1 10-JAN-1986; SHINTO CORP OS Bacillus thuringiensis PN JP 1986005098-A/1 PD 10-JAN-1986 PF 21-FEB-1985 JP 1985031647 PR 22-FEB-1984 US 84 582506, 22-JAN-1985 US 85 693556 PI AAAN EMU UOBUFTIRDO, TOOMASU JIET POROSUKU PC COTK15/04, A01N63/02, C07H21/04, C12N15/00//C12P21/00, (C12N15/00, PC	
COMMENT		

Cl2R1.19);
CC strandedness: Single;
CC topology: Linear;
CC hypothetical: No;
CC anti-sense: No;
CC *source: strain=Igraelensis ONR-60A;
CC *source: clone=BTA4-1;
FH Key Location/Qualifiers
FH 5'UTR <1..925
FT CDS 926..2971
FT FT /product='BRI endotoxin'
FT mat_peptide 926..2968
FT FT /product='BRI endotoxin'
FT 3'UTR 2972..3756.
Location/Qualifiers
source 1..3756
/organism='Bacillus thuringiensis'
/mol_type='genomic DNA'
/db_xref='taxon:1428'
ORIGIN
Query Match 8.1%; Score 178.2; DB 6; Length 3756;
Best Local Similarity 50.0%; Pred. No. 4e-15;
Matches 1108; Conservative 0; Mismatches 923; Indels 186; Gaps 19;
QY 1 ATGAATCAAAATTAAGATATAGCAATATGAAATTTATGATTCGCATACCTCACTTAT 60
DB 941 ATGATTCATATCAAAATTAAGATATAGCAATATGAAATTTATGATTCGCATACCTCACTTAT 994
QY 61 TTTCCGAACAAACAGTATGATTTCTGATACCTTACACAAATATCCAAATCAACCA 120
DB 995 GGTTTTGCAAGTCTAATTAATCTATGATATCCATTAAGCAAAATACCAAAATCAACCA 1054
QY 121 TTACAAAACAAATTAAGATATGATTCGCTCAATATGTCGAAAGGAATATACAAATATGCT 180
DB 1055 CTGAAAAACAGATTTCAAGATTTGCTCAATGTCGTCAGATATCAACAAATATGCT 1114
QY 181 GATAATTCGAGACATTTGCTGATGCTGATACAAATGCTGACGATTAAGTACGATTAAT 240
DB 1115 AATATGCGGGGAATTTTGCTAGTCTGAAATATGTTGAGGTTAGTGACGATTAATTT 1174
QY 241 GTATCCGGTACTCTGTGACCGGATAGGTGGGCTCACTTCTATATCCGACCGATAGGA 300
DB 1175 GTAGTAGAATCTATGTTAGAGCTTTTGTCTGCC-----CT 1210
QY 301 ATAAATAGGTCTATATATATCTTTTGGTACCCATACAGTCTTTTGGCCGCGGA 360
DB 1211 GTCTTAGCTGAGATATATATCTTTTGGACCTTTGTCGACCTTTTGGCAAGATCT 1270
QY 361 GAACAAACAAACAGTATGACACAAATTTATTAATGAGGAATTTTGTGATACA 420
DB 1271 GACCTG---CAATAGTTTGGAGGATTTGTTAAACATCGAGAGGACCTATACAGAA 1327
QY 421 CCGTTTACGAAAGCATTAACAGCTTAAGTTACAACTTTTGAAGGATTTTGAACAATA 480
DB 1328 ATAGATAAAAACTATATTAATGTAATCTT-----CTATCGTAACCTCATATAAAAT 1381
QY 481 TTACAAAGCTATATATAGACATTAAGATGAGAAATTAATAAGCTCAAGCTCT 540
DB 1382 CAATCTGTAATATATCAAGAAATTTTGCATTAATGAGGACGACGATACACGCTAAT 1441
QY 541 GGATTTACACATCATTCAGCATTAACACAGCTGCTTACTTTTAAATATGATTTGAG 600
DB 1442 GCTAAGAGAGTA-CATGATCTCTTAACTAGAACCTAATATATGATTAAGATTTGA 1500
QY 601 AATGTTCAATGATTTTATTCGAGAAATACCTGTTTCAACTTGAAACTTATTAAGCG 660
DB 1501 TATGTTAAAA-----ATAATGCTAGCTATCGATA 1531
QY 661 CTATTTACTACGATTTATGCGAAGCTGTAATTTTCAATTAATTAATTAACAAGGT 720
DB 1532 CCAACATCTCCCTGCAATATGCAAAATAGCTACTTGCACTTGAATTTATTAACATGCT 1591

QY 721 GCTGAATGGCTGATGATGAAATGCAATATACATCTTCACAATTTGAACCTAATGCT 780
DB 1592 GCTACCTATATACATATATATGCTGCAAAATCAAGGTATTAATCCAGTACTTCAATTT-- 1649
QY 781 GGAACATCAATGACTATTTTAACTTTTAAAGAAATATACCTAATATATGATATAT 840
DB 1650 ----CATCTAATTAATATCAAGGGCTATTTTAAACGTAAATATCAAGATATATGACTAT 1705
QY 841 TGTCAATATACCTATAGAACAGACCTAATAAATCTTATAGACGAAACCAATATAGAAATGG 900
DB 1706 TGTATACAAAGTACATGACAGACCTAATATGATTTGAATCTAATATCTAACGCAACTGG 1765
QY 901 AGTATATTTAATAGATATGCAAGATATATGACCAATTAAGTATTTAGATACATCTCAAA 960
DB 1766 AATATGATATATCTTACCGTTTGAATATGACTTAAGTATGATATGATTTGCTATTT 1825
QY 961 TTTTCTTATATGATATTAATAAGATATAGATTTCAATAGAGAAATAGAAATTAAGGC 1020
DB 1826 TTTCCAAATTAAGACCAAGAAATATCC-----AATAGGA 1861
QY 1021 ATTAAGATGAACCTACAAAGAAATTTATACAACTGAATTAATTTTGAATCGTCTTCT 1080
DB 1862 GTTAAATCTGACTATACAGAGATTTATACGAATGTTA----- 1902
QY 1081 CAATTAAGCTTCAACCAATCTAGCTAGATGGAATATATTAACAAGTGAAGTTT 1140
DB 1903 ----TCAGATACATTTAGAACCATTAACAGAACTAGAAATGATTAACAGAAATCT 1957
QY 1141 AATATTTTCAATTTTATAGAACTTTATTTTATACAGAAATATAGAAATTTGCGAAT 1200
DB 1958 ACATTTATTTCTTGATTAACCAAGGGCGTTTATACAGAAATTC--TCGACATTT 2014
QY 1201 CGTTTATGTTGTAATTTCTAATCGTATGACCTATATATAGCATATATATCTGAAT 1260
DB 2015 CTGATACCTATATATATTTTCTTTTACAGTACAGATAGGCTTTACATATCATAT 2074
QY 1261 TTATATGAGAAAGAACAGGTTCAACCAACAAACAAATTAAGACATTTGAATCTTAT 1320
DB 2075 GATGATGCAACATTAATCTGGCGAGCGGTTCAATGAAATATTTATCTCAAGACATCC 2134
QY 1321 AAGTTTCAATTTGATCTGATATGACATCACTCTGTTCCCTATTTCAACCACTTT 1380
DB 2135 AAGATATTTCTCTTTTATTAAGAAACAACTATATATAGATGAAATTTG--CAGACATA 2192
QY 1381 ATAAATATCAAAATGAACTTATTTAATGAGCTCATCTAACAACACATCAATATTTCA 1440
DB 2193 GAGATCTCAGATATATATATGAATGATATTTTGGAAATAGCAGTGAATATTT- 2251
QY 1441 GCAGAGGGCTTTATCTAATTAATCAAAACAACTTTTTCATTTTCAATTTCTTGAAGAAAAA 1500
DB 2252 ----CGATATTCATCAATTCACAAATAGAAAAATATATTAAGAACT----- 2296
QY 1501 GATCGCAATCAATGATTAATCCAGGTTTTCACCAACTTTAATATCTAATGCTATTT 1560
DB 2297 ----GATCTTATATATGATTCAAAACAAACATGAGAAAAATGAAATATGTCATCT 2350
QY 1561 TTATCCATTTTCTATATTTATCTTATTTCTATGATGATTTGATTAAGCTAATATTTA 1620
DB 2351 CTATCGTATATTAATACTGATATTA-----ATATTTCAAGATTTAGGAAGAA 2401
QY 1621 GATACAGGTATTTAGATGACACACAGTATGTTGATATATATATGCAATATTCAGAT 1680
DB 2402 AGAAGATGTCATTTAGTTGGAACACATACATGATTTGATTTCCAAATATACATGATTTA 2461
QY 1681 AAAATATTAATGATATCCAGCAATCAAGATTAACATCTGATATACAACTCTAAGTA 1740
DB 2462 GATACATCAACCAAAATCCAGCTTAATAAACTTTAAGGATTTGTAATTTCAAAATTT 2521
QY 1741 ATTAAGACCTGGTATACAGAGAACTGGTTATTTATTAACAAGTCAAGGCGCTTTA 1800
DB 2522 GTGAAGGCTCGGTCAACAGATGAATACTTGGTATTTCTTAAGATATGATGAATTTT 2581

QY 1801 GAAATTCATGTGAACTCTTAATTTACACATCTTATTTCAATTGACCTTGCATATGCT 1860
DB 2582 AGAGTTAGATTTTAAATAATGTTCTCAACAAATCAAGTACGTA---TTGGTATGCT 2638
QY 1861 ACAATGTTGCTGGAATACCTCTTCAATATATCTTACAAATCCAGGATATAGA 1920
DB 2639 ACTAATGCTCAAGAACACAGTA--TTCTTAAACCGAATAGATATCTAATGTTGAGC 2696
QY 1921 ATACACCTCAACGATCAACAAACACTTTTCTGTGTAACAAATTAATTAATTAATAC 1980
DB 2687 TCCCTAGTACCACTTCCGCCCAAAAC-----CCAAATGCTACGATTTTACATAT 2746
QY 1981 GGAGATTTGGGTATTTCCAAATTTCCAAATGTAAGT---AACATTAACCTTTAAATCGAA 2036
DB 2747 GCAATTTTGGATATGTAACATTTTCAAGAACAGTTTCAAAATTAATAAATTTGAAGAGAA 2806
QY 2037 CATACATTTATATTTA-----ATGTCGAGATGTAATCAATTTCAATTTTATCAT 2088
DB 2807 GACACTTTATTAATGACCTTATATGTTACACCAAAATCAATTCATATTAATATATAT 2866
QY 2089 GATAAAATTTGAATTAATACCAATTAATCTCTATGACCAAAATAGAGAAACAAAA 2148
DB 2867 GACAAATTTGAATTAATTAATCAATCTCAATCTGATTAATTAATTAACAGAGAGAAAT 2926
QY 2149 TTAGAACTATCCAAACAAATAAATAATTTTTCACAAATCAATCAAAACACTT 2205
DB 2927 ATAGAAAAACACGAAATATGTAATGATTTATTTGTTAATTAACAAAGTTCTT 2983

RESULT 15
LOCUS 103578 1993 bp ss-DNA linear PAT 21-MAY-1993
DEFINITION Sequence 1 from Patent US 4652628.
ACCESSION 103578
VERSION 103578.1 GI:268658
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 1993)
AUTHORS Walfield,A.M. and Pollock,T.J.
TITLE Methods and compositions for expression of BTI endotoxin
JOURNAL Patent: US 4652628-A 1 24-MAR-1987;
Syntro Corporation; San Diego, CA
FEATURES
source location/Qualifiers
1..1993
/organism="unknown"
/mol_type="unassigned DNA"

ORIGIN
Query Match 7.8%; Score 171.6; DB 6; Length 1993;
Best Local Similarity 53.4%; Pred. No. 4.1e-14;
Matches 571; Conservative 0; Mismatches 424; Indels 75; Gaps 7;

QY 1 ATGAATCAAAATTAAGATATTAACAATATGAAATTAATTTGATTCGATACCTCACTTAT 60
DB 941 ATGAATCAATTAAGATATTAACAATATGAAATTAATTTGATTCGATACCTCACTTAT 994
QY 61 TTTCCGAACAGAAACAGTAATGATTTCTAGATACCTTTACACAAATATCAATCAACCA 120
DB 995 GGTTTTACAGAGTCAATTAATCAATTTCTAGATATTCATTAAGCAAAATACCAACCA 1054
QY 121 TTACAAACACAAATTAAGAGTGTCAATATGTCAAAGGAAATCAAAATATGAT 180
DB 1055 CTGAAAAACAGAAATTAAGAGTGTCAATATGTCAAAGGAAATCAAAATATGAT 1114
QY 181 GATAATTTGAGACATTTGCTAGTCTGATACAAATGCTGACAGTTAGTGAAGTAT 240
DB 1115 AATAATGCGGGAATTTGCTAGTCTGATACAAATGCTGACAGTTAGTGAAGTAT 1174
QY 241 GTATTCGGTACTCTGTGTGCGGATATAGTGGGCTCACTTCTATATCCGACCATAGA 300
DB 1175 GTAGTAGAACATGTTAGAGCTTTTGTGCTGCTG----- 1211

QY 301 ATAAATAGTGTATTAATTAATATCTTTTGGTACCCCTAATCACTGCTTTTGGCCCGGGA 360
DB 1212 -TCTTAGTGCAGGATATATATCTTTGGACCTTTGTTGCCGATCTTTTGGCAAGATCT 1270
QY 361 GAACAGACAAACAGATGAGACAAATTTATTAATGGAAGAAATTTTGTGATACA 420
DB 1271 GACCTTG---AAATGTTTGGCAGATTTTGAACATCGAGGAGGCCATATCAAGAA 1327
QY 421 CCGTTACAGAACATTAACAGTAAAGTTACAACTTTAGAGAGTTTGAACAAATA 480
DB 1328 ATGATTAACAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1381
QY 481 TTACAAAGCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 540
DB 1382 CACTGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1441
QY 541 GGATTTACCAACATCAATCAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 600
DB 1442 GCTAAAGACAGTA-CATGATCTCTTTAATCACTTGAACCTTAATTAATTAATTAATTA 1500
QY 601 AATGTTACAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 660
DB 1501 TATGTTAAAA-----ATAATGCTAGCTATCGAATA 1531
QY 661 CTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 720
DB 1532 CCAACATCTCTCGATATGACAAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1591
QY 721 GCTGAATGCTGATGAATGGAATGCAATTAATTAATTAATTAATTAATTAATTAATTAAT 780
DB 1592 GCTACCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1648
QY 781 GGAACATCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 840
DB 1649 ---TCATCTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1705
QY 841 TGTCAAAATCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 900
DB 1706 TGTATCAAAAGTATCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1765
QY 901 AGTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 960
DB 1766 AATATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1825
QY 961 TTTTCTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1020
DB 1826 TTTTCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1885
QY 1021 ATTAAGATGAATCAACAGAAATTTATCAATGAAATTAATTAATTAATTAATTAAT 1070
DB 1886 GTTATTAAGATGTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1935

Search completed: December 11, 2005, 17:39:58
Job time : 11160 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OW protein - nucleic search, using frame_plus.p2n model

Run on: January 20, 2006, 17:46:37 ; Search time 300 Seconds
(without alignments)
2019.469 Million cell updates/sec

Title: US-10-783-417-2

Perfect score: 3869

Sequence: 1 NM00NDNNEXEIIIDSHTSPTY.....KLETIQKINTFFVNTKTL 735

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 6049916 seqs, 412136615 residues

Total number of hits satisfying chosen parameters: 12099832

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-Processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODE=frame+ p2n.model -DEV=x1h
-Q/cgn2_1/USPTO.spool/US10783417/runat_20012006_095012_23947/app_query.fasta_1.903
-DB=Published Applications NA New -QMT=fastlap -SUFFIX=p2n.rnpbn -MINMATCH=0.1
-LOOPEXT=0 -LOOPEXT=0 -UNITS=bites -START=1 -END=1 -MATRIX=blosum62
-TRANS-human0.cdi -LIST=45 -DOCLIGN=200 -THR SCORE=pct -THR MAX=100
-THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pct -NORM=exc -HEAPSIZ=500 -MINLEN=0
-MAXLEN=200000000 -USER=US10783417 @CGN_1_121 @runat_20012006_095012_23947
-NCPU=6 -ICPU=3 -NO MMAP -NEG SCORES=0 -WAIT -DSBLOCK=100 -DEV TIMEOUT=120
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications NA New:*

1: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
2: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
3: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
4: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
5: /cgn2_6/ptodata/1/pubpna/US05_NEW_PUB.seq:*
6: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
7: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq1:*
8: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq2:*
9: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq3:*
10: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq3:*
11: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	703	18.2	3633	8 US-11-058-727-3	Sequence 3, Appl1
2	703	18.2	3633	8 US-11-108-389-3	Sequence 3, Appl1
3	703	18.2	6613	8 US-11-058-727-18	Sequence 18, Appl1
4	703	18.2	6613	8 US-11-108-389-18	Sequence 18, Appl1
5	689.5	17.8	4188	10 US-11-091-643-5	Sequence 5, Appl1
6	678.5	17.5	2022	8 US-11-058-727-13	Sequence 13, Appl1
7	678.5	17.5	2022	8 US-11-108-389-13	Sequence 13, Appl1

8	675	17.4	3621	8 US-11-058-727-1	Sequence 1, Appl1
9	675	17.4	3621	8 US-11-108-389-1	Sequence 1, Appl1
10	675	17.4	4874	8 US-11-058-727-17	Sequence 17, Appl1
11	675	17.4	4874	8 US-11-108-389-17	Sequence 17, Appl1
12	669	17.3	4359	10 US-11-091-643-3	Sequence 3, Appl1
13	660	17.1	1959	8 US-11-192-801-3	Sequence 73, Appl1
14	657.5	17.0	2025	8 US-11-058-727-73	Sequence 73, Appl1
15	657.5	17.0	2025	8 US-11-058-727-79	Sequence 79, Appl1
16	657.5	17.0	2025	8 US-11-108-389-79	Sequence 79, Appl1
17	657.5	17.0	2025	8 US-11-108-389-79	Sequence 79, Appl1
18	657	17.0	2022	8 US-11-058-727-81	Sequence 81, Appl1
19	657	17.0	2022	8 US-11-108-389-81	Sequence 81, Appl1
20	656.5	17.0	2025	8 US-11-058-727-41	Sequence 41, Appl1
21	656.5	17.0	2025	8 US-11-058-727-47	Sequence 47, Appl1
22	656.5	17.0	2025	8 US-11-108-389-47	Sequence 47, Appl1
23	656.5	17.0	2025	8 US-11-108-389-47	Sequence 47, Appl1
24	656	17.0	1962	8 US-11-192-801-5	Sequence 5, Appl1
25	656	17.0	2022	8 US-11-058-727-49	Sequence 49, Appl1
26	656	17.0	2022	8 US-11-108-389-49	Sequence 49, Appl1
27	654.5	16.9	2022	8 US-11-058-727-69	Sequence 69, Appl1
28	654.5	16.9	2022	8 US-11-108-389-69	Sequence 69, Appl1
29	654	16.9	1959	8 US-11-192-801-1	Sequence 1, Appl1
30	653.5	16.9	2022	8 US-11-058-727-33	Sequence 33, Appl1
31	653.5	16.9	2022	8 US-11-058-727-67	Sequence 67, Appl1
32	653.5	16.9	2022	8 US-11-108-389-33	Sequence 33, Appl1
33	653.5	16.9	2022	8 US-11-108-389-67	Sequence 67, Appl1
34	653	16.9	2022	8 US-11-058-727-75	Sequence 75, Appl1
35	653	16.9	2022	8 US-11-108-389-75	Sequence 75, Appl1
36	652.5	16.9	2022	8 US-11-058-727-61	Sequence 61, Appl1
37	652.5	16.9	2022	8 US-11-058-727-63	Sequence 63, Appl1
38	652.5	16.9	2022	8 US-11-058-727-65	Sequence 65, Appl1
39	652.5	16.9	2022	8 US-11-108-389-65	Sequence 65, Appl1
40	652.5	16.9	2022	8 US-11-108-389-63	Sequence 63, Appl1
41	652.5	16.9	2022	8 US-11-108-389-65	Sequence 65, Appl1
42	652.5	16.9	2025	8 US-11-058-727-77	Sequence 77, Appl1
43	652.5	16.9	2025	8 US-11-108-389-77	Sequence 77, Appl1
44	652	16.9	2022	8 US-11-058-727-43	Sequence 43, Appl1
45	652	16.9	2022	8 US-11-108-389-43	Sequence 43, Appl1

ALIGNMENTS

RESULT 1
US-11-058-727-3
; Sequence 3, Application US/11058727
; Publication No. US20050261483A1
; GENERAL INFORMATION:
; APPLICANT: Andre R. Abad
; APPLICANT: Ronald D. Flanagan
; APPLICANT: Rafael Herrmann
; APPLICANT: Theodore W. Kahn
; APPLICANT: Albert L. Lu
; APPLICANT: Billy Fred McCutchen
; APPLICANT: James K. Presnall
; APPLICANT: James F. H. Wong
; APPLICANT: Cao-Guo Yu
; TITLE OF INVENTION: Genes Encoding Proteins With Pesticidal
; FILE OF INVENTION: Activity
; FILE REFERENCE: 35718/287809
; CURRENT APPLICATION NUMBER: US/11/058,727
; CURRENT FILING DATE: 2005-02-15
; PRIOR APPLICATION NUMBER: 60/391,786
; PRIOR FILING DATE: 2002-06-26
; PRIOR APPLICATION NUMBER: 60/460,787
; PRIOR FILING DATE: 2003-04-04
; PRIOR APPLICATION NUMBER: 10/606,320
; PRIOR FILING DATE: 2003-06-25
; NUMBER OF SEQ ID NOS: 134
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 3633
; TYPE: DNA
; ORGANISM: Bacillus thuringiensis

Qy	599	-----A	rgleuGluileThrCysgluThrProknserrThrglnserTyrrheilearg	616
Db	1705	CATATGCGCTGTTAAATTTCA-----	-----GACATTAACAAAAGATATATGATAGAGG	1752
Qy	617	LeuArgTyralaThrAsmGlyAlaGlyAsnThrLeuProdenileSerLeuThrilePro		636
Db	1753	ATTGGTATGTGCTTCC-----GCTAATATATCGAATTTATATTAATCCCTTCGAGAA		1806
Qy	637	GlyValileGlyileProProGlnArgLeuAsnThrPheSerGlyThrAsnTyrsn		656
Db	1807	AACCTTAATCT-----CACGCTCAAAAACCTAATGATAGAGGTGAAGCTTTAACTATATAT		1863
Qy	657	AsnLeuGlnTyrrGlyAspPheGlyTyrrheGlnPheProSerThr-----ValThr		673
Db	1864	AAATTTAATATATGCGCACTTTGCCCTTATTAATTTACGACACCGAACCTTTGATTACT		1923
Qy	674	LeuProLeuAsnArgAsnileProPheilePheAsnArgAlaAspValSerAsnSerile		693
Db	1924	CTAGGGGCT-----ATATTGAGCGGAAGACTTTCTTGGAATTGAA		1965
Qy	694	LeuileileAspLysileGlnPheileProileThrSerSerMetHisGlnAsnArgGlu		713
Db	1966	GCTTATATAGACCGAATCGAATTTATCCAGTAGATGAGACATAT-----GAAGCG		2016
Qy	714	LysGlnLysLeuGlnThrIleGlnThrLysileAsnThrPhePheThrAsn		730
Db	2017	GAAACAGATTTAGAAAGCAGCAGAAAGAACAGATGAAATCGCTTTGTTTACCAAT		2067

```

RESULT 2
US-11-108-389-3
; Sequence 3, Application US/11108389
; Publication No. US20050261188A1
; GENERAL INFORMATION:
; APPLICANT: Andre R. Abad
; APPLICANT: Ronald D. Flannagan
; APPLICANT: Rafael Herrmann
; APPLICANT: Theodore W. Kahn
; APPLICANT: Albert L. Lu
; APPLICANT: Billy Fred McCutchen
; APPLICANT: James K. Prensall
; APPLICANT: James F.H. Wong
; APPLICANT: Cao-Guo Yu
; TITLE OF INVENTION: Genes Encoding Proteins With Pesticidal
; Filer REFERENCE: 35718/291049
; CURRENT APPLICATION NUMBER: US/11/108,389
; CURRENT FILING DATE: 2005-04-18
; PRIOR APPLICATION NUMBER: 60/391,786
; PRIOR FILING DATE: 2002-06-26
; PRIOR APPLICATION NUMBER: 60/460,787
; PRIOR FILING DATE: 2003-04-04
; PRIOR APPLICATION NUMBER: 10/606,320
; PRIOR FILING DATE: 2003-06-25
; NUMBER OF SEQ ID NOS: 134
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 3633
; TYPE: DNA
; ORGANISM: Bacillus thuringiensis
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(3633)
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (0)..(0)
; OTHER INFORMATION: Cg1218-2
US-11-108-389-3

Alignment Scores:
Pred. No.: 3,82e-68
Score: 703.00
Percent Similarity: 44.5%

Length: 3633
Matches: 219
Conservative: 127

```

Best Local Similarity:	28.2%	Mismatches:	289
Query Match:	18.2%	Indels:	14
DB:	8	Gaps:	32

US-10-783-417-2 (1-735) x US-11-108-389-3 (1-3633)

QY 5 AenAapAenAenGluTyrGluIleIleAapSerHisThrSerProTyrPheProAenArg 24
 |||||
Db 10 AATATATCAAAATGATATGTAATTTATGATGGACACTTCT-----ACTTCT 57

QY 25 AenSerAenAapSer---ArgTyrProTyrThrAenAenProAenGlnProAenGlnAen 43
 |||||
Db 58 GTATCAATGATCTTCAACAGATACCTTTTGGCAATGAGCCAAACAATGCGTACAAAT 117

QY 44 ThrAenTyrLysGluTyrLeuAenMetCysGlnGlyAen---ThrGlnTyrGlyAenAen 62
 |||||
Db 118 ATGATATTAAAGATTATTTAAATAATGTCGTGGCGGAATGCTAGTCAATATCCCTGGTTCA 177

QY 63 PheGluThrPheAleSerAlaAepThrIleAlaAlaValSerAlaGlyThrIleValSer 82
 |||||
Db 178 CCTGAAGTACTCTTGACGACAAAGAT-----GCAGCTAAAGCCCAATGATGATGA 231

QY 83 GlyThrLeuLeuAlaGlyTyrIleGlyLeuThrSerIleSerGlyProIleGlyIleIle 102
 |||||
Db 232 GGTAAATTAATCACTACAGGTTTAAAGG-----GTCCCAATTGTT 267

QY 103 GlyAlaIleIleIleSerPheGlyThrLeuIleThrValPheTyrProAlaGlyGln 122
 |||||
Db 268 GGGCCGAGATGAGAGCTTATATCTCAACTATATGATATCTGTGGCTTCAGGCGCA--- 324

QY 123 AepTyrThrValTyrThrGlnPheIleLysMetGlyGluIlePheValAepThrProLeu 142
 |||||
Db 325 --AAGAGTCATGAGAGATTTTATTTGGAACAAAGTAAAGAACTCAATATCAAAAATA 381

QY 143 ThrGluSerIleLysGluLeuLysLeuGlnThrLeuGluGlyPheArgGlnIleLeuGln 162
 |||||
Db 382 GCAGAAATATGCAGAGAAATTAAGCCGCTTTCGGAATTGAAAGATTAAGTAAATATTA 441

QY 163 SerTyrAenThrIleLeuAepAepTyrPargLysLeuLysArgLeuGlnAlaProGlyLeu 182
 |||||
Db 442 TTATATCTAACTGCGCTTGAAAGAAAGAAAGAA----- 474

QY 183 ProProSerSerAlaLeuGlnGlnAlaAlaLeuThrLeuLysIleAepPheGluAenVal 202
 |||||
Db 475 AATCCAAATGGTCTA-----AGACGCTTACAGATGTGCGAAATCGATTGAAATCTCG 528

QY 203 HisAenAapPheIleArgGluIleProGlyPheGlnLeuGluThrTyrIleThrLeuLeu 222
 |||||
Db 529 GATAGTTATTTACCGCAATACATGCGCATCTTTTCCAGTACAAATTTTGAAGTACATTC 588

QY 223 LeuProIleTyrAlaGlnAlaAlaAenPheHisLeuAenLeuLeuGlnGlnGlyAlaGlu 242
 |||||
Db 589 CTTACAGATATATACACAGGAGCCAACTTCATTTACTGTATTAAAGACAGCCTTCAATT 648

QY 243 LeuAlaAepGluTyrPheAlaAepIleHisProSerGlnIleGluProAenAlaGlyThr 262
 |||||
Db 649 TTTGGAGAGAGATGGGAGATG-----CTACACACACT 681

QY 263 SerAepAapTyrTyrLysLeuLeuLysGluAenIleProLysTyrSerAenTyrCysAla 282
 |||||
Db 682 ATTATATACATNTATGATCGTCAATATGAACCTTACGCAAGATATTTCTGATCACTGTGA 741

QY 283 AenThrTyrArgThrGlyLeuLysAenLeuAepGluProAenMetLysTyrPheIle 302
 |||||
Db 742 AAGTGGTATGAAACGTTTACCAAAATTTAAAGGACAGAGCGCTAAACAATGGGTGCAC 801

QY 303 PheAenAapTyrArgArgTyrMetThrIleThrValLeuAepThrIleSerGlnPheSer 322
 |||||
Db 802 TATATCACAATTCGATGAGAAATATGACATCGACGTTTATGATGCTTCCATTAATCCCA 861

QY 323 LeuTyrAepIleLysArgTyrArgAepSerIleGlyGlyIleGluValLysGlyIleLys 342
 |||||
Db 862 AATTATGACACGACGCTAC-----CCATGGAACGAAGAACA----- 900

Db 1371 AAGGATTATAAGATTATTTAAATAATGTCGGCGGAAATGCTAGTAATACCTCGGTCA 1430
 QY 63 PheGluThrPheAlaSerAlaAspThrIleAlaIleValSerAlaGlyThrIleValSer 82
 Db 1431 CCTGAAGTACTGTTGATCGGACAAAGAT-----GCGAGCTAAGCGCCCAATTGATTAATAGA 1484
 QY 83 GlyThrLeuLeuAlaGlyIleGlyIleLeuThrSerIleSerGlyProIleGlyIleIle 102
 Db 1485 GGTAAATTACTATACAGGTTTAGG-----GTCCATTGTT 1520
 QY 103 GlyAlaIleIleIleSerPheGlyThrLeuIleThrValPheThrProAlaGlyGluGln 122
 Db 1521 GGGCCGATGATGAGCTTATTAATCACTTATGATATCTGTGGCTTCAGGGCA--- 1577
 QY 123 AspIleThrValIleThrGlnPheIleIleIleMetGlyGluIlePheValAspThrProLeu 142
 Db 1578 ---AAGATCAATGGAGATTTTATGACAAAGTGAAGAACTCAATCAATAAATA 1634
 QY 143 ThrGluSerIleGlyGlnLeuIleGlnThrLeuGluGlyPheArgGlnIleLeuGln 162
 Db 1635 GCGAATATGCAAGAAATTAAGCGCTTGGAAATTTGAAGATTAAGTAATATTAACCA 1694
 QY 163 SerTyrAsnThrAlaLeuAspAspArgTyrGlyIleLeuIleValArgLeuGlnAlaProGlyLeu 182
 Db 1695 TTATATCTAATCGCTGTAAGAAATGAAAGAA----- 1727
 QY 183 ProProSerSerAlaLeuGlnGlnAlaIleLeuThrLeuIleValLeuArgPheGluAsnVal 202
 Db 1728 AATCCAAATGGTTCA-----AAGACCTTAACGAGATGCGAAATGGAATTTGAAATCCG 1781
 QY 203 HisAsnAspPheIleArgGluIleProGlyPheGlnLeuGlnIleThrTyrIleThrLeuLeu 222
 Db 1782 GATAGTTTATTTACGCAATACATGCACTTTTCGATGACAAATTTGAAGTACCATTC 1841
 QY 223 LeuProIleTyrAlaGlnAlaIleAsnPheHisLeuAsnLeuLeuGlnGlnIleAlaGln 242
 Db 1842 CTTCAGTATATACACAGGACGCAACCTTCATTTCGTATTAAGACGCTTCGAAT 1901
 QY 243 LeuAlaAspGluIleTyrAsnAlaAspIleHisProSerGlnIleGluProAsnAlaGlyThr 262
 Db 1902 TTTGGAGAAATGCGGATG-----TCTACAAACCACT 1934
 QY 263 SerAspAspTyrTyrIleLeuLeuIleValAsnIleProIleThrSerAsnTyrCysAla 282
 Db 1935 ATTAATTAATCTATTAATATGTCATAATGAACTTACGCAAAATATTTCTGATCTGTGA 1994
 QY 283 AsnThrTyrArgThrGlyLeuIleValAsnLeuArgAspGluProAsnMetIleTyrSerIle 302
 Db 1995 AAGTGTATGAAACTGTTTAGCAAAATTTAAAGGACAGCGCTTAACATGCGTGCAC 2054
 QY 303 PheAsnAspTyrArgArgTyrMetThrIleThrValLeuAspThrIleSerGlnPheSer 322
 Db 2055 TATAACCAATTCGATGAGAAATGACATGACGTTTGTAGATGTGTGCATTAATCTCA 2114
 QY 323 LeuTyrAspIleIleArgTyrArgAspSerIleGlyIleGluValIleGlyIleIle 342
 Db 2115 AATTATGACACAGCGATAC-----CCAATGGAACGAAAGCA----- 2153
 QY 343 AsnGluLeuThrArgGluIleIleTyrThrGlnIleAsnPheAspArgLeuProGlnLeu 362
 Db 2154 ---CAACTAACAAGAGAAATATAC-----GATCCACTGGCGCGCGGTA 2195
 QY 363 ArgValGln-----ProAsnLeuAlaThrMetGluTyr 373
 Db 2196 AACGTCTCTCAATGTGTCCTGTATGACAAAGCACTTCCTTCGAGATGATAGATACA 2255
 QY 374 AsnLeuThrArgAlaSerPheIleLeuPheSerPheLeuGlnGlnIleIlePheTyrThr 393
 Db 2256 TCCGTATTTGACACACC---CATGTATTTGATTATTAACGGAAGCTCACAGTGTATACA 2312
 QY 394 GluAsnThrAsnPheGluAsn-----ArgLeuValGlyIleSerAsnArg 408
 Db 2313 CAATCAAGAGCACTTTCTCCGCTGCTATATAAGACATTTGGCTGTCTCAATAAAGC 2372

QY 409 AspAlaProThrTyrSerAsnThrIleThrGluThrLeuTyrGlyAlaArgThrGlySer 428
 Db 2373 TATCATCGGATTTTATGATATATTAATAAAGATATAGAACTATCAATAATCTA 2432
 QY 429 ProThrThrIleThrIleArgProPheGluSerTyr-----Ile 441
 Db 2433 CACAGACTAGTACCTT---GATTTACGAATTAATGATTTTACAAAGCTTATCAAA 2489
 QY 442 ValSerIleValThrAspArgGlnSerProProValSerProIleGlnProHisPheIle 461
 Db 2490 GATGGCGTCCCTGATATGTTGTTTCTCGTTATACATAATA-----TTTTTGA 2543
 QY 462 IleAsnGlnIleGluLeuIleValAsnGlySerSerAsnAsn-----ThrLeuIle 478
 Db 2544 ATGCGAAGATGAGATTTTCAATGTAACCAATGGAATTAATCAAGAAAGCGTTAAAG 2603
 QY 479 TyrSer-----AlaGlyGlySerLeuSerAsnTyrGlnAsnThr 491
 Db 2604 TATATCCGCTTCCAAAGATTAATTAAGCGGG----- 2636
 QY 492 ThrPhePheGlnPheProArgIleIleValAspCysAsnLeuValIleAspProGlyCysSer 511
 Db 2637 -----ACAAGAGATTCGGAATTAATTAATCACTCAGAACTTCA 2675
 QY 512 -----ProAsnPheAsnAsnTyrSerHisIleLeuSerHisPhe---SerLeuPheThr 528
 Db 2676 GATCAACCAATTAATGATCATATACCATATGATTAATGTCATATACCAAGTATCCCGCG 2735
 QY 529 TyrSerTyrValIleGlyLeuGlnIleGlnIleLeuAspThrGlyValIleGlyTyrThr 548
 Db 2736 ACGGGTCAACATACCGGATTA-----GTACCTGATTTTCTTGAC 2777
 QY 549 HisSerSerValAspArgTyrAsnAlaIleSerAspIleIleIleThrMetIleProAla 568
 Db 2778 CATCGAGTGCAGATCTTAATATGACATTCATTCAGATTAATTAATCAATCCGCGTC 2837
 QY 565 IleIleGlyAsnAsnLeu-----AspThrAsnSerIleValIleGlu 582
 Db 2838 GTAAAGTTTGTGATTTGGCTCCCTCTATACAGAGCGCCAAATTAATCCGTTGATCG 2897
 QY 583 GlyProGlyHisIleThrGlyIleValAsnLeuValTyrLeuGlnSerGlnIle 598
 Db 2898 GGTCTCGATTTACAGGGGGGAGATTAATAAGTAATGAAGTGAATTAATTAATCA 2957
 QY 599 -----ArgLeuGluIleThrCysGluThrProAsnSerThrGlnSerTyrPheIleArg 616
 Db 2958 CATATGCGGTTAATAATTCA-----GACATTAACAAAGATATATGATGAGCG 3005
 QY 617 LeuAlaGlyTyrAlaThrAsnGlyAlaGlyAsnThrLeuProAsnIleSerLeuThrIlePro 636
 Db 3006 ATTCGGTATGCTTC-----GCTAATAATCTGAATTTTATTAATATCTTCTGAAGA 3059
 QY 637 GlyValIleGlyIleProGlnArgLeuAsnAsnThrPheSerGlyThrAsnTyrAsn 656
 Db 3060 AACGTTAATCT---CACCTCAAAAACATAAGATAGAGGAGAACTTAACATTAAT 3116
 QY 657 AsnLeuGlnTyrGlyAspPheGlyTyrPheGlnPheProSerThr-----ValThr 673
 Db 3117 AATTATTAATTAATGCACTTGGCCCCCTTAATTAATTAATTAACAAACCAACTTCATTACT 3176
 QY 674 LeuProLeuAsnArgAsnIleProPheIlePheAsnArgAlaAspValSerAsnSerIle 693
 Db 3177 CTAGGGGCT-----ATAATTAAGCGGAAGACTTCTTGGAATTTGA 3218
 QY 694 LeuIleIleAspIleIleGluPheIleProIleThrSerSerMetHisGlnIleAsnArgIle 713
 Db 3219 GCTTATATGACAGCAATCAATTAATTAATCCCAAGTATGAACAATAT-----GAAGCG 3269
 QY 714 IyGlnIleValLeuGluIleThrIleGlnThrIleValAsnThrPhePheThrAsn 730
 Db 3270 GAACAAAGATTAAGAGCAAGCAAGAAAGCAAGTAAGTCCCTGTATTGCAAT 3320

```

RESULT 4
US-11-108-389-18
; Sequence 18, Application US/11108389
; Publication No. US2005026118A1
; GENERAL INFORMATION:
; APPLICANT: Andre R. Abad
; APPLICANT: Ronald D. Flammagan
; APPLICANT: Rafael Herrmann
; APPLICANT: Theodore W. Kahn
; APPLICANT: Albert L. Lu
; APPLICANT: Billy Fred McCutchen
; APPLICANT: James K. Presnall
; APPLICANT: James F. H. Wong
; APPLICANT: Cao-Guo Yu
; TITLE OF INVENTION: Genes Encoding Proteins with Pesticidal
; FILE REFERENCE: 35718/291049
; CURRENT APPLICATION NUMBER: US/11/108,389
; PRIOR FILING DATE: 2005-04-18
; PRIOR APPLICATION NUMBER: 60/391,786
; PRIOR FILING DATE: 2002-06-26
; PRIOR APPLICATION NUMBER: 60/460,787
; PRIOR FILING DATE: 2003-04-04
; PRIOR APPLICATION NUMBER: 10/606,320
; NUMBER OF SEQ ID NOS: 134
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 18
; LENGTH: 6613
; TYPE: DNA
; ORGANISM: Bacillus thuringiensis
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (0)...(0)
; OTHER INFORMATION: Genomic Cyt1218-2
US-11-108-389-18

Alignment Scores:
Pct. No.: 1.03e-67 Length: 6613
Score: 703.00 Matches: 219
Percent Similarity: 44.5% Conservative: 127
Best Local Similarity: 28.2% Mismatches: 289
Query Match: 18.2% Indels: 142
Gaps: 32
DB: 8

US-10-783-417-2 (1-735) x US-11-108-389-18 (1-6613)

QY 5 AenAaPaaenAaenGluTyrGluIleIleAaPaaenHleThrSerProTyrPheProAaenArg 24
DB 1263 AATTAATCAAAATGAATATGAATTAATTAATGATGACCACTTCT-----ACTTCT 1310
QY 25 AaenSerAaenAaPaaen---ArgTyrProTyrThraAaenProAaenGlnProleuGlnAaen 43
DB 1311 GATATCAATGATTTCTAACAAGATACCTTTTGGCAATGACCAACAATGCGCTACAAAT 1370
QY 44 ThraenTyrLysGluTyrLeuAaenMetCysGlnGlyAaen---ThrGlnTyrGlyAaen 62
DB 1371 ATGCAATTATTAAGATATTATTAATAATGTCGCGGGAATGCTAAGTAATACCTGCTCA 1430
QY 63 PheGluTyrPheAaenSerAaenAaPaaenHleAaIaIaValSerAaIaGlyThrIleValSer 82
DB 1431 CCTGAAGTACTGTTAGCGGACAAGAT-----GCAGCTAAGCGCGCAATTGATATAGTA 1484
QY 83 GlyThrLeuLeuAaIaGlyIleGlyGlyLeuThrSerIleSerGlyProIleGlyIleIle 102
DB 1485 GGTAAATTAATCTATCGAGTTAGG-----GTCCCATTTGTT 1520
QY 103 GlyAaIaIeIleIleSerPheGlyThrLeuIleThraValPheTyrProAaIaGlyGln 122
DB 1521 GGGCGCATGTGAGCTTATATCTCAACTATTAATTTCTGTGGCTTCAAGGGA--- 1577
QY 123 AaPaaenThraValTyrThrGlnPheIleLysMetGlyGluIlePheValAaPaaenThraProleu 142
DB 1577 AaPaaenThraValTyrThrGlnPheIleLysMetGlyGluIlePheValAaPaaenThraProleu 142

```

```

DB 1578 ---AAGACTCAATGGAGATTTTATGGAAACAATGAGAACTCATTAATCAAAAAATA 1634
QY 143 ThrGluSerIleLysGluIleuLysLeuGluIleThrLeuGluGlyPheArgGlnIleuGln 162
DB 1635 GCAGAAATATGCAAGAAATTAAGCGCTTCCGAAATTAAGAAAGATTAGTAATTAATCAACA 1694
QY 163 SerTyrAaenThraValLeuAaPaaenThraArgLysLeuLysAaGlyGlnAaIaProGlyLeu 182
DB 1695 TTATATCTAATCGCGCTTGAAGAAATGAGAAAGAA----- 1727
QY 183 ProPaaenSerAaIaLeuGlnGlnAaIaIaLeuThrLeuLysIleArgPheGluAaenVal 202
DB 1728 AATCCAAATGGTTC-----AGAGCTTACAGATGCGCAATGCGAATTTGTAATTCCTCG 1781
QY 203 HisAaenAaPaaenPheIleArgGluIleProGlyPheGlnLeuGluTyrThrLysThrLeu 222
DB 1782 GATAGTTATTTAGCAATATACCAATGCACTTTTCCAGAGACAAATTTGAATATACCATTC 1841
QY 223 LeuProIleTyrAaIaGlnAaIaAaenPheHisLeuAaenLeuGlnGlnGlyAaIaGlu 242
DB 1842 CTTACAGATATATACACAGGACAGCCAACTTCATTTACTGTTATTAAGACCGCTTCAATT 1901
QY 243 LeuAaenPaaenGluTyrAaenAaIaAaPaaenHleHisProSerGlnIleGluProAaenAaIaGlyThr 262
DB 1902 TTTGAGAGAAATGGGAGATGC-----CTTAACAACACT 1934
QY 263 SerAaenAaPaaenTyrTyrLysLeuLeuLysGluAaenIleProLysTyrSerAaenTyrCysAaIa 282
DB 1935 AATTAATCAATTAATGATGATGCTCAATTAAGAACTTATGCAAGATATTTGATACACTGTGTA 1994
QY 283 AaenThraTyrArgThrGlyLeuLysAaenLeuAaPaaenGluProAaenMetLysTyrPaaenIle 302
DB 1995 AAGGTGATGAACACTGTTTAGCAAAATTAAGACGACGCGCTAAACAATGGCTCGAC 2054
QY 303 PheAaenAaPaaenTyrArgArgTyrMetThrIleThrValLeuAaPaaenThraIleSerGlnPheSer 322
DB 2055 TATTAACCAATTCCTGTAAGAAATGACCTGACGCTTTAGATGTTGCTGATTAATTCCTCA 2114
QY 323 LeuTyrAaPaaenIleLysArgTyrArgAaPaaenIleGlyGlyIleGluValLysGlyIleLys 342
DB 2115 AATTATGACACACGACGATAC-----CCAATGGAACGAACAA----- 2153
QY 343 AaenGluLeuThraArgGluIleTyrThrThrGluIleAaenPheAaPaaenProGluLeu 362
DB 2154 ---CAACTAACAAGGAAATATATACA-----GATCCACTGGCGCGCGTA 2195
QY 363 ArgValGln-----ProAaenLeuAaIaThrMetGluTyr 373
DB 2196 AACGTGCTTCAATTTGTTCTCTGTAATGACAAAGACCTTCTTGGAGTGATGAAATCA 2255
QY 374 AaenLeuThraArgAaIaSerPheLysLeuPheSerPheLeuGlnPheIlePheTyrThr 393
DB 2256 TCCGTATTAACCAACCC---CATGTATTTGATTAATTAACGGGACCTCAAGGTATATCA 2312
QY 394 GluAaenThraAaenPheGlyAaen-----ArgLeuValGlyIleSerAaenArg 408
DB 2313 CAATCAAGAAGCATTTCTTCGCTGCTATATTAACATGCGCTGTCATCAATTAAGC 2372
QY 409 AaPaaenAaProThraTyrSerAaenThraIleThraGluThraLeuTyrGlyGluAaArgThrLys 428
DB 2373 TATCATCGCAATTTTATGATATATTAATTAACGAATGATGACATTAATCAAAATCTTA 2432
QY 429 ProThraThraLysThraIleArgProPheGluSerTyr-----Lys 441
DB 2433 CACAGCACTAGACTTT---GATTTACGAATTAATGATATTTACAGACGTTATCAAAA 2489
QY 442 ValSerIleValThraAaPaaenGlnSerProProValSerProIleGlnProHisPheIle 461
DB 2490 GATGCGGCTTCCTGATATATGTTTCTCTGTTATAGTATATA-----TTTTTGGA 2543
QY 462 IleAaenGlnIleGluLeuTyrLeuAaenGlySerAaenAaen-----ThraLeuLys 478
DB 2544 ATGCCAAGATCGAATTTTTCATGTGTAACCAATTAATTAATCAAGAAAGCGTTAAAG 2603

```

```

QY 479 TyrSer-----AlaGlyGlySerLeuSerAsnTyrGlnAsnThr 491
DB 2604 TATTAATCCGGTTTCCAAAGATATATATAGCGGG----- 2636
QY 492 ThrPhePheGlnPheProArgLysLysAspCysAsnLeuValIleAspProGlyCysSer 511
DB 2637 -----ACAGAGATTCGGAAATTGAAATTACCTCCAGAAACTTCA 2675
QY 512 -----ProAsnPheAsnAsnTyrSerHisIleLeuSerHisPhe---SerLeuPheThr 528
DB 2676 GATCAACCAAAATTATGATCATATAGCCATATATGTCATATACCAAGATATTCGCCGG 2735
QY 529 TysSerTyrValIleGlyLeuGlnLeuGlnIleLeuAspThrGlyValLeuGlyTyrThr 548
DB 2736 ACGGCTTCAACTACCGGATTA-----GTACCTGATATTTCTTGGACA 2777
QY 549 HisSerSerValAspArgTyrAsnAlaIleSerAspLysIleIleThrMetIleProAla 568
DB 2778 CATCGGAGTGCAGATCTTATTAATGATGATTCAGATTAATATCTCAGATTCGGGTC 2837
QY 569 IleLysGlyAsnAsnLeu-----AspThrAsnSerLysValIleGlu 582
DB 2838 GTAAAGGTTCTGATTGGCTCCCTCTATACAGAGGCGCAATATATACGTTGTATCG 2897
QY 583 GlyProGlyHisThrGlyLysAsnLeuValTyrLeuGlnSerGlnGly----- 598
DB 2898 GGTCTCTGATTATACAGGGGGGGGATATAAAGATATAGAAATGAGATTAATTATATCA 2957
QY 599 -----ArgLeuGlnIleThrCysGlnThrProAsnSerThrGlnSerTyrPheIleArg 616
DB 2958 CATATGCGGTATAAATTCA-----GACATTAACAAGAAATATATGATGAGG 3005
QY 617 LeuArgTyrAlaThrAsnGlyAlaGlyAsnThrLeuProAsnIleSerLeuThrIlePro 636
DB 3006 ATTCGGATGCTTCC-----GCTATATATACGATTTTATATTAATCTCTCGAAGA 3059
QY 637 GlyValIleGlyIleProProGlnArgLeuAsnAsnThrPheSerGlyThrAsnTyrAsn 656
DB 3060 AACGTTAATCT---CACGCTCAAAAACATATGAAATAGCGAGACCTTAAACATATAT 3116
QY 657 AsnLeuGlnTyrGlyAspPheGlyTyrPheGlnPheProSerThr-----ValThr 673
DB 3117 AATTTAATATATGCGACTTGGCCCCCTATTAATTTAGACAACGACCTTCTACTACT 3176
QY 674 LeuProLeuAsnArgAsnIleProPheIlePheAsnArgAlaAspValSerAsnSerIle 693
DB 3177 CTAGGGGGCT-----ATATTGAAACCGAAGACCTTCTTGGATTGAA 3218
QY 694 LeuIleIleAspLysIleGlnPheIleProIleThrSerSerMetHisGlnAsnArgGlu 713
DB 3219 GCTTATATAGACCAATCGAATTTATCCAGTAGATGACACATAT-----GAAGCG 3269
QY 714 LysGlnLysLeuGlnIleGlnThrLysIleAsnThrPhePheThrAsn 730
DB 3270 GAACAAGATTGAAAGCAGCAAGAAAGACGTAAGCTTGTATGCAAT 3320

RESULT 5
US-11-091-643-5
; Sequence 5, Application US//11091643
; Publication No. US20050246789A1
; GENERAL INFORMATION:
; APPLICANT: TANAKA, Masao
; APPLICANT: YOKOYAMA, Tomoko
; APPLICANT: AOYAGI, Morichi
; APPLICANT: HASEGAWA, Makoto
; APPLICANT: EHARA, Gaku
; APPLICANT: KIMURA, Masaharu
; APPLICANT: NISHIHASHI, Hideji
; TITLE OF INVENTION: Polypeptide having larvae growth inhibiting or
; TITLE OF INVENTION: Insecticidal effect on scarabaeidae insects and
; FILE REFERENCE: OP1335

```

```

; CURRENT APPLICATION NUMBER: US/11/091,643
; CURRENT FILING DATE: 2005-03-29
; PRIOR APPLICATION NUMBER: JP 2001-115754
; PRIOR FILING DATE: 2001-04-13
; PRIOR APPLICATION NUMBER: JP 2001-203463
; PRIOR FILING DATE: 2001-07-04
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5
; LENGTH: 4188
; TYPE: DNA
; ORGANISM: Bacillus popilliae
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(4158)
US-11-091-643-5

Alignment Scores:
Pred. No.: 1,65e-66 Length: 4188
Score: 689.50 Matches: 239
Percent Similarity: 45.1% Conservative: 113
Best Local Similarity: 30.6% Mismatches: 286
Query Match: 17.8% Indels: 143
DB: 10 Gaps: 38

US-10-783-417-2 (1-735) x US-11-091-643-5 (1-4188)

QY 1 MetAsnGln-----AsnAsnAspAsn-----AsnGluTyr 10
DB 55 ATGATATCATATATCATATACCAACCAATGTAACCAACCAAGTGAATGCAATGATG 114
QY 11 GlnIleIle-----AspSerHisThrSerProTyrPheProAsnArgAsnSerAsnAsp 28
DB 115 CAATTTATCAACCTTCAAGTACGCTTTACTTTACAGTCCCAAC----- 159
QY 29 SerArgTyrProTyrThrAsnAsnProAsnGlnProLeuGlnAsnThrAsnTyrLysGlu 48
DB 160 ---AGTATCCGATATGCTACGAGATCCCAATGTCATACAGAGAGTATGAAATTAAT 216
QY 49 TrpLeuAsnMetCysGlnGlyAsnThrGlnTyrGlyAspAsnPheGlnThrPheAlaSer 68
DB 217 TGGTTGGATATGCTGTAGT-----GTAGGCTGACGAT-----ACA 252
QY 69 AlaAspThrIleAlaAlaValSerAlaGlyThrIleValSerGlyThrLeuLeuAlaGly 88
DB 253 CGAATCCCGAAGCTCGGGTATCTGCACAAAGTTCTATTGCACTTCTT----- 303
QY 89 IleGlyGlyLeuThrSer---IleSerGlyProIleGlyIle-----IleGlyAlaIle 105
DB 304 -----GGTATACACGACACATATCTGGCGCTTACGATATTCGGTGTAGCGCAAGCC 357
QY 106 IleIleSerPheGlyThrLeuIleThrValPheTyrProAlaGlyGlnGlnAspLysThr 125
DB 358 ATTTGAATTTTGGCGCACTATTTGATTTGTGTGGCTGTAGCGGCTGATCCA----- 411
QY 126 ValTyrThrGlnPheIleLysMetGlyGlnIlePheValAspThrProLeuThrGlnSer 145
DB 412 ---TGGTATATATTTATGATCATCATAGAACTCATTAATTTCAAAAATACAGAGACT 468
QY 146 IleLysGlnLeuLysLeuGlnThrLeuGlnGlyPheArgGlnIleLeuGlnSerTyrAsn 165
DB 469 GTAAATAATAGAGCAATTAACAAGATTAGACGTTTATAGCAATGCTCTACGCTATATCA 528
QY 166 ThrAlaLeuAspAspTyrArgLysLysLysAsnValSerGlnAlaProGlyLeuProSer 185
DB 529 AAGCTTTTCAAGATAGCGCA-----CAACATCT----- 558
QY 186 SerAlaLeuGlnGlnAlaAlaLeuThrLeuLysIleArgPheGlnAsnValHisAsnAsp 205
DB 559 ---ACTCTTGAAATCACTGCACTACGTGACAGATGATTTTCTTAATGTAATAATTT 615
QY 206 PheIleArgGlnIleProGlyPheGlnLeuGlnGlnThrTyrLysThrLeuLeuProIle 225

```


Oy	-----AGLeagGluIleThhCyselunThrProAensertInserTlyPhelellyg	616
Db	1705 CATATGCCTTTAAATTCA-----GACATTACAAGAATATATGRTAGCG	1753C
Oy	617 LeuArgTyralathraengLYAlaGLyaenthreuProaenlleSerleuthrilePro	636
Db	1753 ATTGGGTATGCTTCC-----GCTATATATCTGAATTTATATAACCCTTCGAAGAA	1806C
Oy	637 GlyValIlelegIleProProdlnargLeuaenanthruPheSergLythrAsnTyrsan	656
Db	1807 AACGTTAATCT---CACGCTCAAAAACTAGATTAAGCGAAGACTTTAACTATAT	1865C
Oy	657 AsnLeugIntYrgLYaspPhnegLYtyrPheglInpheProserThr-----Valthr	673
Db	1864 AAATTTAATTATMGCACTTGCCGCCCTTAAATTATACAGAACCACTTCACTTACT	1923C
Oy	674 LeuProLeuaasnArgsnlleProPheIlePheasnaArgalaAspValserFamsberille	693
Db	1924 CTAGGGGCT-----ATATTGAAAGCGGAAGACTTCTTGGAATTGAA	1965C
Oy	694 LeuIlelleaplyalleglnPheIleProille	704
Db	1966 GCTTATATAGACCGAATCGAATTTATCCACAGTA	1998

OY	363	ArgValGln-----	-----ProAsnLeuAlaThiMetGluTyr	373
Db	943	AACGTCCTCAATTGGTCTCGTAGACAAAGACACCTCTTTCGGAGTAGATGAATCA	1002	
OY	374	AsnLeuThrArgAlaSerPheIleuPheSerPheIleuGluGlnPheIlePheTyrThr	393	
Db	1003	TCCGTTATTCAACCAACC-----CATGATTATTTGATTATATACGGAGATCAAGTGTATATCA	1059	
OY	394	GluAsnThrAsnPheGlyAsn-----	-----ArgLeuValGlyIleSerAsnArg	408
Db	1060	CAATCAAGAGACATTTCTTCGCTCGTATATTAAGACATTTGGCGTGCATCAATATAGC	1119	
OY	409	AspAlaProThrTyrSerAsnThrIleThrGluThrLeuTyrGlyIuArgThrGlySer	428	
Db	1120	TATCATCGGATTTTATGATATATATATATATATTAACAGATGTATGAACTTAATCAAAATCTA	1179	
OY	429	ProThrThrIlyThrIleArgProPheGluSerTyr-----	-----Lys	441
Db	1180	CACAGACACTAGAACCTTT--GATTTACGAATTATGATATTTACAGACGTATACAAA	1236	
OY	442	ValSerIleValThrAspArgGlnSerProValSerProIleGlnProHisPheIle	461	
Db	1237	GATGGGGGCTCCCTGATATATGTTTTCCTGGTATATACGTATATA-----TTTTTGA	1290	
OY	462	IleAsnGlnIleGluLeuTyrLeuAsnGlySerSerAsnAsn-----	-----ThrLeuLys	478
Db	1291	ATGCGAGAGTCGAGTTTTCATGTGTAAACCAATGTGAATATATACAGAAACGTTAAAG	1350	
OY	479	TyIser-----	-----AlaGlyGlySerLeuSerAsnTyrGlnAsnThr	491
Db	1351	TATATATCCGGTTTCCAAAGATATATATACGGGG-----	-----	1383
OY	492	ThrPhePheGlnPheProArgIlyLeuSerCysAsnLeuValIleAspProGlyCysSer	511	
Db	1384	-----ACAAGAGATTGGGAATTAGAAATTACCTCCAGAAACTTCA	1422	
OY	512	-----ProAsnPheAsnAsnTyrSerHisIleLeuSerHisPhe--	-----SerLeuPheThr	528
Db	1423	GATCAACCAAAATTATAGATCATATATGCCATATGATTATATCATACAAAGATATCCCGCG	1482	
OY	529	TyrSerTyrValIleGlyLeuGlnLeuGlnIleLeuAspThrGlyValLeuGlyTyrThr	548	
Db	1483	ACGGGTCACTACCGGATTA-----	-----GTAACGTATTTTCTTGACA	1524
OY	549	HisSerSerValAspArgTyrAsnAlaIleSerAspLysIleIleThrMetIleProAla	568	
Db	1525	CATCGAGATGCGCGACTTATATATACGATTCATTCAGATAAATATATCTCAGATTCGCGTC	1584	
OY	569	IleIysGlyAsnAsnLeu-----	-----AspThrAsnSerLysValIleGlu	582
Db	1585	GTAAGGTTCCTGATTGGTCCCTCTATTAACAGGAGGCGCAAAATTAATATACCGTGTATCG	1644	
OY	583	GlyProGlyHisThrGlyGlyAsnLeuValTyrLeuGlnSerGlnGly-----	-----	598
Db	1645	GCTCTCGATTATACAGGGGGGGGATATATAAAGATATAGAAATGAGATATTATATCA	1704	
OY	599	-----ArgLeuGluIleThrCysGluThrProAsnSerThrGlnSerTyrPheIleArg	616	
Db	1705	CATATGCGCTGTAAATTTCA-----	-----GACATTAACAAAGATATATGATATGAG	1752
OY	617	LeuArgTyrAlaThrAsnGlyValaGlyAsnThrIleuProAsnIleSerLeuThrIlePro	636	
Db	1753	ATTCGGTATGCTTC-----GCTATATATATCTGAATTTATATTAATATCTTCGAGAA	1806	
OY	637	GlyValIleGlyIleProProGlnIleArgLeuAsnAsnThrPheSerGlyThrAsnTyrAsn	656	
Db	1807	AACGTTAATCT--CACGCTCAAAAACCTATGATATAGAGGTGAAGCTTTAATCAATATAT	1863	
OY	657	AsnLeuGlnTyrGlyAspPheGlyTyrPheGlnPheProSerThr-----	-----ValThr	673
Db	1864	AAATTTATATATGCACTTTGCGCCCTATTAATTTACGACACCGAAACCTTGATTAATCT	1923	
OY	674	LeuProLeuAsnArgAsnIleProPheIlePheAsnArgAlaAspValSerAsnSerIle	693	

```

Db      1294 CTAAGGGCT-----ATATTTAAGCGAAGACTTTCTTGAAATTGAA 1965
QY      694 Leuileleaplysliegupheileproile 704
      1966 GCTTATATATAGACCGAATGCAATTATTCACAGTA 1998

RESULT 8
US-11-058-727-1
; Sequence 1, Application US/11058727
; Publication No. US20050261483A1
; GENERAL INFORMATION:
; APPLICANT: Andre R. Abad
; APPLICANT: Ronald D. Flannagan
; APPLICANT: Rafael Herrmann
; APPLICANT: Theodore W. Kahn
; APPLICANT: Albert L. Lu
; APPLICANT: Billy Fred McCutchen
; APPLICANT: James K. Prensall
; APPLICANT: James F.H. Wong
; APPLICANT: Cao-Guo Yu
; TITLE OF INVENTION: Genes Encoding Proteins with Pesticidal
; TITLE OF INVENTION: Activity
; FILE REFERENCE: 35718/287809
; CURRENT APPLICATION NUMBER: US/11/058,727
; PRIOR FILING DATE: 2005-02-15
; PRIOR APPLICATION NUMBER: 60/391,786
; PRIOR FILING DATE: 2002-06-26
; PRIOR APPLICATION NUMBER: 60/460,787
; PRIOR FILING DATE: 2003-04-04
; PRIOR APPLICATION NUMBER: 10/606,320
; PRIOR FILING DATE: 2003-06-25
; NUMBER OF SEQ ID NOS: 134
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 3621
; TYPE: DNA
; ORGANISM: Bacillus thuringiensis
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(3621)
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (0)...(0)
; OTHER INFORMATION: Cpy1218-1
US-11-058-727-1

Alignment Scores:
Pred. No.:      5.78e-65      Length:      3621
Score:          675.00      Matches:      216
Percent Similarity: 44.5%      Conservative: 125
Best Local Similarity: 28.2%      Mismatches:   300
Query Match:    17.4%      Indels:       126
DB:             8      Gaps:         32

US-10-783-417-2 (1-735) x US-11-058-727-1 (1-3621)

QY      5 Aenbapbanaanglutyrgluillelaspserhishrserprotyrphneproasnarg 24
      10 AATATCAAAATGATATGAAATATATGATGCGACACCTTCT-----ACTTCT 57
QY      25 Aamseraenapssar---ArgtyrProtyrThrashanpnoaenglProleuglInasn 43
      58 GTATTCAGATGATCTTAACAGATACCTCTTTCGGAATGACCCAAATGCGCTACAAAT 117
Db      44 ThraenTyrlvgIutrlpLeuAsnmetCysgInglYasn--ThrgInTyrgIyAspAsn 62
      118 ATGGATTATTAAGAATTAATTAAATATGCTGCGGGAATGCTAGTGAATACCTGAGTTCA 177
QY      63 PheglutlrhrpealaserlaspThrillealalavalserlagnlyThrillevalser 82
      178 CCGTAAGTACTTGTTAGGCGACAAGAT-----GCACCTAAGGCGCGCAATTGATGATGTA 231

```


/ APPLICANT: Ronald D. Flanagan
 / APPLICANT: Rafael Herrmann
 / APPLICANT: Theodore W. Kahn
 / APPLICANT: Albert L. Lu
 / APPLICANT: Billy Fred McClutchen
 / APPLICANT: James K. Presnall
 / APPLICANT: James F.H. Wong
 / APPLICANT: Cao-Guo Yu
 / TITLE OF INVENTION: Genes Encoding Proteins With Pesticidal
 / TITLE OF INVENTION: Activity
 / FILE REFERENCE: 35718/291049
 / CURRENT APPLICATION NUMBER: US/11/108,389
 / CURRENT FILING DATE: 2005-04-18
 / PRIOR APPLICATION NUMBER: 60/391,786
 / PRIOR FILING DATE: 2002-06-26
 / PRIOR APPLICATION NUMBER: 60/460,787
 / PRIOR FILING DATE: 2003-04-04
 / PRIOR APPLICATION NUMBER: 10/606,320
 / PRIOR FILING DATE: 2003-06-25
 / NUMBER OF SEQ ID NOS: 134
 / SOFTWARE: FastSeq for Windows Version 4.0
 / SEQ ID NO 1
 / LENGTH: 3621
 / TYPE: DNA
 / ORGANISM: Bacillus thuringiensis
 / FEATURE:
 / NAME/KEY: CDS
 / LOCATION: (1)...(3621)
 / FEATURE:
 / NAME/KEY: misc feature
 / LOCATION: (0)...(0)
 / OTHER INFORMATION: Cyt1218-1
 / US-11-108-389-1

Alignment Scores:

Pred. No.: 5,78e-65 Length: 3621
 Score: 675.00 Matches: 216
 Percent Similarity: 44.5% Conservative: 125
 Best Local Similarity: 28.2% Mismatches: 300
 Query Match: 17.4% Indels: 126
 DB: 8 Gaps: 32

US-10-783-417-2 (1-735) x US-11-108-389-1 (1-3621)

QY 5 AaNaPaSaNaSgLuTrGluLeIleAaPSeRHiSThSeRProTyRPhRoBaNaTy 24
 DB 10 AATATCAAAATGAATATGAAATTAATAGATGCGACACTTCT-----ACTTCT 57
 QY 25 AaNaPaSaNaSgLuTrGluLeIleAaPSeRHiSThSeRProTyRPhRoBaNaTy 43
 DB 58 GATCCAAATGATTTCTAAGATACCTCTTGGCAATGAGCCCAAAATGCGCTCAAAAT 117
 QY 44 ThRaSeTyLySgLuTrPLeuSaNeMeCySgInGlyAaSn---ThRgInTyRgLyAaPaN 62
 DB 118 ATGCAATTAATGAATATTAATTAATATGTCGCGGGAATGCTAGTAATATCCCTGATCA 177
 QY 63 PhEgLuThPhEaLaSeRaLlaSPThRlIeLaLaIaVaSeRaLaGlyThRlIeLaSeR 82
 DB 178 CCTGAAGTACTGTGATGCGGACAAAGAT-----GCAGCTAAGCGCCCAATGATATAGTA 231
 QY 83 GlyThRleuLaLaGlyLeGlyLeuThRSeRlIeSeRgLyProIleGlyIleIe 102
 DB 232 GGTAAATATCTATGAGGTTTAAAGG-----GTCCCATTTGTT 267
 QY 103 GLYAlaIleIleIleSeRPhEgLyThRleuLeThRValPhETRProLaGlyGluIn 122
 DB 268 GGGCCGATAGTGAAGCTTATATCACTTATGATATCTGCGGCTTCAAGGGA--- 324
 QY 123 AaPlyThRValTrPrThRgInPhEIlIeLyMeGlyGluIlePhEValaSPThRProLeu 142
 DB 325 ---AAGAGTCAATGAGAAATTTTATAGAAACAAGTGAAGAACTCATTAATCAAAATA 381
 QY 143 ThRgUseRlIeLySeRlIeUeLySeuGInThRleuGInGlyPhEaRgGInIleuGIn 162

DB 382 GCAGATATGCAAGAAATTAAGCGCTTCCGAATTAGAAGATAGATATATACCA 441
 QY 163 SeTyRaenThRaIleuAaPaSPThRgLySLeuLyAaRgLeuGInLaProGlyLeu 182
 DB 442 TTATATCTAATCGCCCTTGAAGAAATGGAAGAA----- 474
 QY 183 ProPosSeSeRaLleuGInGInLaIleuThRleuLyIleAaRPhEgLuSaVaL 202
 DB 475 AATCCAAATGCTTCA-----AGAGCTTACAGAGTGTCCGAATTCGATTTGAATCTCG 528
 QY 203 HiSaNaSPPhEIlIeAaRgLyLeProGlyPhEgInLeuGluThRlyThRleuLeu 222
 DB 529 GATAGTTATTAACCAATATATGCACTTTAGAGTGAACAATTTGAAGTACATTC 588
 QY 223 LeuProIleTyRaIleGInLaIaLaSPPhEIlIeLySeuAeNleuGInGInLaGly 242
 DB 589 CTACTGATATATGCAATGCGCAACCTTCATTTAAGTATTAAGCAAGCGCTCAAT 648
 QY 243 LeuAlaSPGluThRPaSaNaLaSPThRlIeSPSeRgInIleGluProBaNaLaGlyThR 262
 DB 649 TTGGAGAAAGATGGGATGC-----TCAACAATCTACT 681
 QY 263 SeRaPaSPTyRlySLeuLeuLySgLuAeNlIeProLyTySeRaNTyCySaLa 282
 DB 682 ATTAATTAATTAATGATGTCGAATGAACCTTACGCAATATCTGATCACTGTGA 741
 QY 283 AaenThTyRaTyThRgLyLeuLySaNeuAaRgLeuProBaNeMeLySPSeRlIe 302
 DB 742 AAGGTATGAACCTGTTTACCAAAATTAAGCAAGCGGCTTAACATGCGGTGAC 801
 QY 303 PhEaNaSPTyRaTyThRgLyMeThRlIeThRValIleuAaSPThRlIeSeRInPSeR 322
 DB 802 TATAACCAATTCGTAAGAAATGAACCTGCGGTTTAAAGTGTGCTCATTAATCCCA 861
 QY 323 LeuTyRaPleLyAaRgTyRaRgAaSPSeRlIeGlyLyIleGlyValIleGlyLyLe 342
 DB 862 AATTATGACACACGACGATC-----CCAATGAAGAAACA----- 900
 QY 343 AaenGluLeuThRaGlyLeuTyThRgLyLeuThRgLyLeuThRgLyLeuThRgLyLeu 362
 DB 901 ---CAACTAACAAGGAAGATATATACA-----GATCCATGGCGCGGTA 942
 QY 363 AaRgValGIn-----ProBaNeLaIaThMeGlyTyR 373
 DB 943 AAGGTCTTCAATGCTTCTGATAGCAAGACCTTCTTGGAGTATAGATCA 1002
 QY 374 AaenLeuThRaGAlaSeRPhElySeRPhESeRPhELeuGluGInPhEIlIePhElyTyR 393
 DB 1003 TCCGTTATTCGACCAACC---CATGATTTGATTTATATACGGGACTCACAGTATACA 1059
 QY 394 GlUAenThRaSPPhEgLyAaSn---AaRgLeuVal-----GlyIleSeR 406
 DB 1060 CAATCAAGAAGCATTTCTCGCTGCTATATAGCAATGCGGCTGATCAAAATPAC 1119
 QY 407 AaNaPaSaPaLaProThTySeRaenThRlIeThRgInThRleuTyRgLyGlyAaRgThR 426
 DB 1120 TACCAATCGTCAAGTGGGTAGTAT-----CTTCAACAATATGATGAACATATCA 1173
 QY 427 GlySeRProThThRlyThRlIeAaRProPhEgIuSeRlyTyRlyAaSeRlIeValThR 446
 DB 1174 AATCTACACAGCATGATACCTT---GATTTACGAATATGATATTAACAAGACTCTA 1230
 QY 447 AaPaRgGInSeRProBaVaSeRProIleGInPro-----HiSPheIle 461
 DB 1231 TCAAGGAATGAGTACTCTGATATATGTTTACCCTGATATACGATATATTTTGA 1290
 QY 462 IleAenGInIleGlyLeuTyRleuAeNlySeRSeRaNaN-----ThRleuLy 478
 DB 1291 ATGCCAAGAGTGAAGTTTTCATGAGTAAACAATGAATATATCAAGAAAGCGTTAAAG 1350
 QY 479 TySeRaLaGlyLySeRleuSeRaNTyRgInAenThThRPhEgInPhEProAaRg 498

Db 1351 TAT-----AATCGATTTCGAAGATATTATGCGCATGACAGA 1389
Qy LysylsAspCyAsnLeuValIleAspProGlyCySerProAsnPhenAsnTyrSer 518
Db 1390 GATTGGAAATTGAAATTAACCTCCAGAACTTGAGTCAACAAATTAATGAGCATATAGC 1449
Qy HselleuSerHisPhe---SerleuPheThrTyrSerTyrValIleGlyleuGlnLeu 537
Db 1450 CATGATATTATGTCATATCAACAGATATCCCGGACGGGTACACCTACCGGATTA----- 1503
Qy GlnIleleuAspThrGlyValleuGlyTyrThrHisSerSerValAspArgTyrAsnAla 557
Db 1504 -----GTACCTGTAATTTCTTGGACACATCGAAGTCAGATTAACAATACA 1551
Qy HleSerAspLysIleIleThrMetIleProAlaIleLys---GlyAsnAsnLeuAspThr 576
Db 1552 ATATATTGAGTAATAATACCTCAATTCGGCGGCTTAATGTTGGGATTAATTTACCGTTT 1611
Qy AsnSerLysValIleGlyGlyProGlyHisThrGlyGlyAsnLeuVal---TyrleuGln 595
Db 1612 ---GTTCCAGTGGTAAAGACACAGACATACAGAGGGGATTTATACAGTATATAGA 1668
Qy SerGlnGly-----ArgleuGlnIleThrCyGlyleuThrPro 607
Db 1669 AGTACTGGTCTGTAGGAACTTATTCTAGCTCGATGATGCGCTAGCATTAAGA----- 1722
Qy AsnSerThrGlnSerTyrPheIleArgLeuArgTyrAlaThrAsnGlyValIleGlyAsnThr 627
Db 1723 ---AAGACGGGAATATGCTGTAGACATGAGATATGCTAGATGCA----- 1767
Qy LeuProAsnIleSerleuThrIleProGlyValIleGlyIleProProGlnArgleuAsn 647
Db 1768 -----GATATGTATGTGATGTAAACGATGCTCAGATTCAGATGCCAAAACATGAGAC 1821
Qy AsnThrPheSerGlyThrAsnTyrAsnLeuGlnTyrGlyAspPheGlyTyrPheGln 667
Db 1822 CCAAGT-----GAGGATCTGACATCTTAAACTTAAAGTTGCAAT 1863
Qy PheProSerThrValThrleuProleuAsnArgAsnIleProPheIlePheAsnArgAla 687
Db 1864 GGTATCACAACATTAATTATGACACATGCTGCTAGCATGGAACATATTAATGCT 1923
Qy AspValSerAsnSer-----IleleuIleIleAspLysIleGlnPheIlePro 703
Db 1924 GAAGACCCCTAATTCACATTAATCTGGTATAGTTTACGTGACCGAATCGAATTCACCA 1983
Qy 704 IleThrSerSerMetHisGlnAsnArgGlyLysGlnLysleuGlnThrIleGlnThrLys 723
Db 1984 GTAGATGAGACATAT-----GAAAGCGGACAAAGATTATGAAAGCAGCAAGAAACACA 2034
Qy 724 IleAsnThrPhePheThrAsn 730
Db 2035 GTGAATGCTTGTTTACGAAT 2055

RESULT 10
US-11-058-727-17
; Sequence 17, Application US/11058727
; Publication No. US20050261483A1
; GENERAL INFORMATION:
; APPLICANT: Andre R. Abad
; APPLICANT: Ronald D. Flannagan
; APPLICANT: Rafael Herrmann
; APPLICANT: Theodore W. Kahn
; APPLICANT: Albert L. Lu
; APPLICANT: Billy Fred McCutchen
; APPLICANT: James K. Presnail
; APPLICANT: James F. H. Wong
; APPLICANT: Cao-Guo Yu
; TITLE OF INVENTION: Genes Encoding Proteins with Pesticidal
; FILE REFERENCE: 35718/287809
; CURRENT APPLICATION NUMBER: US/11/058,727
; CURRENT FILING DATE: 2005-02-15

; PRIOR APPLICATION NUMBER: 60/391,786
; PRIOR FILING DATE: 2002-06-26
; PRIOR APPLICATION NUMBER: 60/460,787
; PRIOR FILING DATE: 2003-04-04
; PRIOR APPLICATION NUMBER: 10/606,320
; PRIOR FILING DATE: 2003-06-25
; NUMBER OF SEQ ID NOS: 134
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17
; LENGTH: 4874
; TYPE: DNA
; ORGANISM: Bacillus thuringiensis
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (0)...(0)
; OTHER INFORMATION: Genomic DNA 1218-1
US-11-058-727-17

Alignment Scores:
Pred. No.: 9,466-65 Length: 4874
Score: 675.00 Matches: 216
Percent Similarity: 44.54 Conservative: 125
Best Local Similarity: 28.24 Mismatches: 300
Query Match: 17.48 Indels: 126
Gaps: 32

US-10-783-417-2 (1-735) x US-11-058-727-17 (1-4874)

Qy 5 AsnAspAsnAsnGluTyrGluIleIleAspSerHisThrSerProTyrPheProAsnArg 24
Db 740 AATAATCAAAATGAATATGAAATTAATGATGACACCTTCT-----ACTTCT 787
Qy 25 AsnSerAsnAspSer---ArgTyrProTyrThrAsnAsnProAsnGlnProleuGlnAsn 43
Db 786 GATATCATGATTTCTACACAGATACCTTTTGGATATGAGCCACAAAGACGCTACAAAT 847
Qy 44 ThrAsnTyrLysGluTyrleuAsnMetCyGlnGlyAsn---ThrGlnTyrGlyAspAsn 62
Db 848 ATGATATTAATAAGATTATTAATAATGCTCGCGGAAATGCTAGTGAATACCTCGTTCA 907
Qy 63 PheGluThrPheAlaSerPheAlaAspThrIleAlaIleValSerAlaGlyThrIleValSer 82
Db 908 CCGTAAGTACTTGTATGAGGACAGAT-----GCAAGCTAAGCGCGCAATGATATAGTA 961
Qy 83 GlyThrLeuLeuAlaGlyIleGlyGlyleuThrSerIleSerGlyProIleGlyIleIle 102
Db 962 GGTAAATTAATCTACATCGCTTTAGG-----GTCCCATTTGTT 997
Qy 103 GlyAlaIleIleIleSerPheGlyThrleuIleThrValPheTyrProAlaGlyGlnGln 122
Db 998 GGGCCGATAGTAGAGCTTATTAATCTCACTTATGATATTCTGTGGCCTTACAGGGGAA--- 1054
Qy 123 AspLysThrValTyrThrGlnPheIleLysMetGlyGluIlePheValAspThrProleu 142
Db 1055 ---AAGAGTCATAGGAAATTTTATGAAACAAGTAGAAGAACTCATTAATCAAAAAATA 1111
Qy 143 ThrGluSerIleLysGlnleuLysleuGlnThrleuGlnGlyPheArgGlnIleleuGln 162
Db 1112 GCAGAAATTCAGAGAAATTAACCGCTTTCGGAATTAAGAAAGATTAGTAATAATTAACCA 1171
Qy 163 SerTyrAsnThrAlaLeuAspAspTyrArgLysleuLysArgleuGlnAlaProGlyleu 182
Db 1172 TTATATCTAATCGCTTGAAGATAGGAAGA----- 1204
Qy 183 ProProSerSerAlaLeuGlnGlnAlaIleuThrleuLysIleArgPheGlnAsnVal 202
Db 1205 AATCCAAATGGTTCA-----AGAGCCTTACAGATGCGGAATCGATTGAATCCCTCG 1258
Qy 203 HsAsnAspPheIleArgGluIleProGlyPheGlnleuGluThrTyrLysThrleuLeu 222
Db 1259 GATAGTTTATTAACCAATATATAGCAATCTTTTAAGTAGACAAATTTTGAATACCATTC 1318
Qy 223 LeuProIleTyrAlaGlnAlaIleAsnPheHisleuAsnLeuGlnGlnIleGlyAlaGlu 242

```

Db      1319 CTTACTGATATGCAATGAGCGACCACTTCACTTATCTGTATTAAGACGGCGCAANT 1378
Qy      243 LeuAlaAspGluTrpAsnAlaAspIleHisProSerGlnIleGluProAsnAlaGlyThr 262
Db      1379 TTTGAGAGAAATGGCGATGG-----TCACCAACTACT 1411
Qy      263 SerAspAspTyrTyrIleuLeuLeuGluAsnIleProIleTyrSerAsnTyrCysAla 282
Db      1412 ATTTAATTAATCTATTATATGCTCAATGAAATCTTACGAGAAATATCTGATCTGTGTA 1471
Qy      283 AsnThrTyrArgThrGlyLeuLysAsnLeuArgAspGluProAsnMetIleTyrSerIle 302
Db      1472 AAGGTGATGAACACTGTTTATAGCAAAATTAAGACAGCGGCTAAACATGGGTGAC 1531
Qy      303 PheAsnAspTyrArgArgTyrMetThrIleThrValLeuAspThrIleSerGlnPheSer 322
Db      1532 TATTAACCAATTCGGTGAAGAAATGACACTGGCGTTTATGATGTGTGCATTATATCCCA 1591
Qy      323 LeuTyrAspIleLysArgTyrArgAspSerIleGlyIleGluValIleGlyIleLys 342
Db      1592 AATTATGACACACGACGTAC-----CCAATGGAACGAAGCA----- 1630
Qy      343 AsnGluLeuThrArgGluIleTyrThrThrGluIleAsnPheAspArgLeuProGlnLeu 362
Db      1631 ---CAACTAACAGAGAAAGTATACA-----GATCCTCGGGCGGGCGTA 1672
Qy      363 ArgValGln-----ProAsnLeuAlaThrMetGluTyr 373
Db      1673 AACGTCTCTTCATTTGCTTCGTGTATGACAAAGCACTTCTTCGAGATGATGAATGA 1732
Qy      374 AsnLeuThrArgAlaSerPheLysLeuPheSerPheLeuGlnGlnIlePheTyrThr 393
Db      1733 TCCGTATTTCGACCAACC---CATGTATTGATTATTAACGGGACTCAACAGTGTATCA 1789
Qy      394 GluAsnThrAsnPheGlyAsn---ArgLeuVal-----GlyIleSer 406
Db      1790 CATCAAGAGACATTTCTCCGCTCCATATTAAGACATTTGGCTGCTATCAAAATTAAC 1849
Qy      407 AsnArgAspAlaProThrTyrSerAsnThrIleThrGluThrLeuTyrGlyArgThr 426
Db      1850 TACATCGTGTACGTAGGGGTACTAAT---CTTCAACAATGATGGAATCAATCA 1903
Qy      427 GlySerProThrThrIleArgProPheGlnSerTyrLysValSerIleValThr 446
Db      1904 AATCTACACAGCACTGTACCTTT---GATTTACCAATTAATTAATTAACAGACTCTA 1960
Qy      447 AspArgGlnSerProThrValSerProIleGlnPro-----HisPheIle 461
Db      1961 TCAAAGATGCACTACTCTTGATATTTGTTTACCTCGTTATACGTATATTTTTCGA 2020
Qy      462 IleAsnGlnIleGluLeuTyrLeuAsnGlySerSerAsnAsn-----ThrLeuLys 478
Db      2021 AAGCCAGAGTCAAGTTTTCATGTGAACCAATTAATTAATTAACGAAGAGCTTAAG 2080
Qy      479 TyrSerAlaGlyLysLeuSerLeuSerAsnTyrGlnAsnThrThrPheGlnPheProGly 498
Db      2081 TAT-----AATCCAGTTTCCAAAGATTAATTAAGCACTACACAA 2119
Qy      499 LysLysAspCysAsnLeuValIleAspProGlyCysSerProAsnAsnAsnTyrSer 518
Db      2120 GATTCGAATTAAGATTAATCTCCAGAAACTCGATCAACCAATTAATTAAGTATTAAG 2179
Qy      519 HisIleLeuSerHisPhe---SerLeuPheThrTyrSerTyrValIleGlyLeuGlnLeu 537
Db      2180 CATGATTAATGTCATATCAACAAGTATTCGCGACGGTAAACACTACCGGATTA----- 2233
Qy      538 GlnIleLeuAspThrGlyValIleGlyTyrThrHisSerSerValAspArgTyrAsnAla 557
Db      2234 -----GTACCTGTATTTCTTGGACACATCGAAGTCAAGATTTAAACAATTAAC 2281
Qy      558 IleSerAspLysIleIleThrMetIleProAlaIleLys---GlyAsnAsnLeuAspThr 576

```

```

Db      2282 ATATATTCAGATTAATCACTCAAAATCCGCGCTTAATGTGGATTAATTACCGTTT 2341
Qy      577 AsnSerLysValIleGlnGlyProGlyHisThrGlyGlyAsnLeuVal---TyrLeuGln 595
Db      2342 ---GTTCCAGGTGTAAAGACACAGACATTAACGAGGGGATTTATTAAGTATTAATGA 2398
Qy      596 SerGlnGly-----ArgLeuGluIleThrCysGluThrPro 607
Db      2399 AGTACTGGTTCGTAGGAACCTTATTTACGCTCGATATAGCGCTACGATTAGA----- 2452
Qy      608 AsnSerThrGlnSerTyrPheIleArgLeuArgTyrAlaIleThrAsnGlyAlaGlyAsnThr 627
Db      2453 ---AAAGCAGGAAATATCGTGTAAAGACTGAGATATGCTACGATGCA----- 2497
Qy      628 LeuProAsnIleSerLeuThrIleProGlyValIleGlyIleProProGlnArgLeuAsn 647
Db      2498 -----GATATTGTATGATGTAAACGATGCTCAGATTCAGATGCAAAACAATGAAC 2551
Qy      648 AsnThrPheSerGlyThrAsnTyrAsnAsnLeuGlnTyrGlyAspPheGlyTyrPheGln 667
Db      2552 CCAGGT-----GAGATCTGACATCTTAAACTTTAAAGTTCAGAT 2593
Qy      668 PheProSerThrValThrLeuProLeuAsnArgAsnIleProPheIlePheAsnArgAla 687
Db      2594 GGTATCAACAATTAATTAGCAACAGATAGTGTGCTGACATTAATTAATTAAGGT 2653
Qy      688 AspValSerAsnSer-----IleLeuIleIleAspLysIleGluPheIlePro 703
Db      2654 GAAGACCTTAATTAACATTAATCTGTATAGTTTACGTGTGACGATCGAATTCATCCCA 2713
Qy      704 IleThrSerSerMetHisGlnAsnArgGlyLysGlnLeuGluThrIleGlnThrLys 723
Db      2714 GTAGATGACATCAT-----GAGCGGACAAAGATTTAGAACACGGAAGAAAGCA 2764
Qy      724 IleAsnThrPhePheThrAsn 730
Db      2765 GTGAATGCTGTGTTACGAAT 2765

RESULT 11
US-11-108-389-17
; Sequence 17, Application US/11108389
; Publication No. US20050261188A1
; GENERAL INFORMATION:
; APPLICANT: Andre R. Abad
; APPLICANT: Ronald D. Flanagan
; APPLICANT: Rafael Herrmann
; APPLICANT: Theodore W. Kahn
; APPLICANT: Albert L. Lu
; APPLICANT: Billy Fred McCutchen
; APPLICANT: James K. Presnail
; APPLICANT: James F. H. Wong
; APPLICANT: Cao-Guo Yu
; TITLE OF INVENTION: Genes Encoding Proteins With Pesticidal
; FILE REFERENCE: 35718/291049
; CURRENT APPLICATION NUMBER: US/11/108,389
; PRIOR FILING DATE: 2003-04-18
; PRIOR APPLICATION NUMBER: 60/391,786
; PRIOR FILING DATE: 2002-06-26
; PRIOR APPLICATION NUMBER: 60/460,787
; PRIOR FILING DATE: 2003-04-04
; PRIOR APPLICATION NUMBER: 10/606,320
; PRIOR FILING DATE: 2003-06-25
; NUMBER OF SEQ ID NOS: 134
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17
; LENGTH: 4874
; TYPR: DNA
; ORGANISM: Bacillus thuringiensis
; FEATURE:
; NAME/KEY: misc. feature
; LOCATION: (0)...(0)
; OTHER INFORMATION: Genomic DNA 1218-1

```

US-11-108-389-17

Alignment Scores:

Pred. No.: 9,466-65 Length: 4874
 Score: 675.00 Matches: 216
 Percent Similarity: 44.5% Conservative: 125
 Best Local Similarity: 28.2% Mismatches: 300
 Query Match: 17.4% Indels: 126
 Ds: 8 Gaps: 32

US-10-783-417-2 (1-735) x US-11-108-389-17 (1-4874)

QY 5 AenAspAsnAsnGluTyrGluIleIleAspSerHisThrSerProTyrPheProAsnArg 24
 DB 740 AATTAACCAAAATGAAATATGAAATATATAGATGCGACACTTCT-----ACTTCT 787
 QY 25 AenSerAsnAspSer---ArgTyrProTyrThrAsnAsnProAsnGlnProLeuGlnAsn 43
 DB 788 GTATCCAAATGATCTTAAACAGATACCTCTTTCGCAATGACCAACAAATGCGCTACAAAT 847
 QY 44 ThrAsnTyrLeuGluTyrPheAsnMetCysGlnGlyAsn---ThrGlnTyrGlyAspAsn 62
 DB 846 ATGGATATATTAAGATATTTAAATAATGTCGCGGAAATGCTAGTAATACCTGCTTCA 907
 QY 63 PheGluThrPheAlaSerAlaAspThrIleAlaIleValSerAlaGlyThrIleValSer 82
 DB 908 CCTGAAGTACTTGTTCGCGACCAAGAT-----GCAGCTAAGCGCCCAATGATATATAGA 961
 QY 83 GlyThrLeuLeuAlaGlyIleGlyGlyLeuThrSerIleSerGlyProIleGlyIle 102
 DB 962 GGTAAATTAATCTATACAGGTTTAGG-----GTCCCAATTTGTT 997
 QY 103 GlyAlaIleIleIleSerPheGlyThrLeuIleThrValPheTyrProAlaGlyIleGln 122
 DB 998 GGGCGCATGAGTACTTATATCTCAACTATGATATCTGCGCTTCAAGGGA--- 1054
 QY 123 AspLysThrValTyrThrGlnPheIleLysMetGlyGlnIlePheValAspThrProLeu 142
 DB 1055 ---AAGAGTCAATGGCAATTTTATATGAAACAAATGAAACCTCATATCAAAAAATTA 1111
 QY 143 ThrGluSerIleLysGlnLeuLysGlnInThrLeuGlnGlyPheArgGlnIleLeuGln 162
 DB 1112 GCAGAAATATGCAAGGAAATAAAGCGCTTTCGCAATTAAGAAAGATAGTAATATTAACAA 1171
 QY 163 SerTyrAsnThrAlaLeuAspAspTyrArgLysLeuLysArgLeuGlnAlaProGlyLeu 182
 DB 1172 TATATCTAAGCTGCGCTTGAAGATGGAAGAA----- 1204
 QY 183 ProProSerSerAlaLeuGlnGlnAlaAlaLeuThrLeuLysIleArgPheGluAsnVal 202
 DB 1205 AATCCAAATGAGTTCA-----AGAGCCTTACGAGATGTCGAAATGATTTGAAATCTCG 1258
 QY 203 HisAsnAspPheIleArgGluIleProGlyPheGlnLeuGluThrTyrIleThrLeuLeu 222
 DB 1259 GATAGTTTATTAACGAAATATATGCAATCTTTTGAAGTACAAATTTTAAAGTACCAATTC 1318
 QY 223 LeuProIleTyrAlaGlnAlaAlaAsnPheHisLeuAsnLeuGlnGlnGlyAlaGln 242
 DB 1319 CTTACTGTAATGCAATGCGACGCAACTTCATTTCTGTATTTAAAGACCGCTCAAT 1378
 QY 243 LeuAlaAspGluTyrAsnAlaAspIleHisProSerGlnIleGluProAsnAlaGlyThr 262
 DB 1379 TTTGGAGAGAAATGGGGATGG-----TCACCAACTACT 1411
 QY 263 SerAspAspTyrTyrIleLeuLeuLysGluAsnIleProLysTyrSerAsnTyrCysAla 282
 DB 1412 ATTAATTAACATATATATGTCGAATGAAACTTACGCAAAATATCTATCACTGTGTGA 1471
 QY 283 AsnThrTyrArgThrGlyLeuLysAsnLeuArgAspGluProAsnMetLysTyrSerIle 302
 DB 1472 AAGGTATAGAAACTGCTTTTGGCAAAATTTAAAGGACAGACGCTTAAACAATGGGTTGAC 1531
 QY 303 PheAsnAspTyrArgArgTyrMetThrIleThrValLeuAspThrIleSerGlnPheSer 322

DB 1532 TATAACCAATTCCTAGTAGAAGAAATGACACTGGCGGTTTATAGATGTTGCAATATTCACA 1591
 QY 323 LeuTyrAspIleLysArgTyrArgAspSerIleGlyGlyIleGluValIleGlyIleLys 342
 DB 1592 AATTATGACACACGACAGTAC-----CCATGGAAACGAAACA----- 1630
 QY 343 AsnGluLeuThrArgGluIleTyrThrThrGluIleAsnPheAspArgLeuProGlnLeu 362
 DB 1631 ---CACTTACACAGGAGAGTATATACA-----GATCCACTGGGCGCGGTA 1672
 QY 363 ArgValGln-----ProAsnLeuAlaThrMetGluTyr 373
 DB 1673 AACGTCCTTCAATGTTCTCTGATAGCAAAAGACCTTCTTTCGAGTGTATAGAAATCA 1732
 QY 374 AsnLeuThrArgAlaSerPheLysPheSerPheLeuGluGlnPheIlePheTyrThr 393
 DB 1733 TCCGTTATTCACACACACC---CATGTATTTGATTTATATACGGGACTCAAGTATATACA 1789
 QY 394 GluAsnThrAsnPheGlyAsn---ArgLeuVal-----GlyIleSer 406
 DB 1790 CAATCAAGACGATTTCTTCGCTGATATATAGACATTTGGGCTGTGCATCAAAATTAAC 1849
 QY 407 AsnArgAspAlaProThrTyrSerAsnThrIleThrGluThrLeuTyrGlyGluArgThr 426
 DB 1850 TACCATCGTGTCAAGTAGGGGTATAT-----CTTCAACAAATGATATGAACTATATCA 1903
 QY 427 GlySerProThrThrLysThrIleArgProPheGluSerTyrLysValSerIleValThr 446
 DB 1904 AATCTACACAGCACTAGTACCTTT---GATTTAGCAATATATGATTTATCAAGACTCTA 1960
 QY 447 AspArgGlnSerProValSerProIleGlnPro-----HisPheIle 461
 DB 1961 TCAAAGATAGCAGTACTCCTGATATATGTTTACCTCGGTATATACGATATATTTTGGCA 2020
 QY 462 IleAsnGlnIleGluLeuTyrLeuAsnGlySerSerAsnAsn-----ThrLeuLys 478
 DB 2021 ATGCCAAGAGTCAAGTATTTCTATGTGTAAACCAATTTGAAATATACCAAGAAAGCGTTAAG 2080
 QY 479 TyrSerAlaGlyLysLeuSerLeuSerAsnTyrGlnAsnThrPhePheGlnPheProArg 498
 DB 2081 TAT-----AATCCAGTTTCCAAAGATATATATACCGATACAA 2119
 QY 499 LysLysAspCysAsnLeuValIleAspProGlyCysSerProAsnPheAsnAsnTyrSer 518
 DB 2120 GATTCGGAATTAAGATATACCTCCAGAACTTCAGATCAACCAATATATAGCATATAC 2179
 QY 519 HisIleLeuSerHisPhe---SerLeuPheThrTyrSerTyrValIleGlyLeuGlnLeu 537
 DB 2180 CATGATTAATGTCATATACCAAGATATTCCTCCGCAAGGGTATACCTACCGATTA----- 2233
 QY 538 GlnIleLeuAspThrGlyValLeuGlyTyrThrHisSerSerValAspArgTyrAsnAla 557
 DB 2234 -----GTACCTGATTTTCTTGACACATTCGAAAGTCAGATTTTAAACAATATACA 2281
 QY 558 IleSerAspLysIleIleThrMetIleProAlaIleLys---GlyAsnAsnLeuAspThr 576
 DB 2282 ATATATTCAGATTAATATCACTCAAAATTCGCGCGGTTTAAATGTTGGGATATATTAACGTTT 2341
 QY 577 AsnSerLysValIleGluGlyProGlyHisThrGlyGlyAsnLeuVal---TyrLeuGln 595
 DB 2342 ---GTTCCAGTGTAAAGGACAGACATACAGAGGGGATTTTATACAGTATATATAGA 2398
 QY 596 SerGlnGly-----ArgLeuGlnIleThrCysGluThrPro 607
 DB 2399 AGTACTGCTTCTGTAAGAACTTATTTCTAGCTCGATATAGCCCTACGATTAAGA----- 2452
 QY 608 AsnSerThrGlnSerTyrPheIleArgLeuArgTyrAlaThrAsnGlyValGlyAsnThr 627
 DB 2453 ---AAAGCAGGAAATATATGTTAGACATGATATGCTACGATGCA----- 2497
 QY 628 LeuProAsnIleSerLeuThrIleProGlyValIleGlyIleProGlnArgLeuAsn 647

Db 2498 -----GATATTGATTCAGATGTAACGATGCTCAGATTCAGATGCCAAAACATGTAAC 2551
 Qy 648 AanthrPhseSerGlyThrAntyRaanaenLeuGlnTyGlyAspPheGlyTyRheGln 667
 Db 2552 CCAAGT-----GAGGANTCTACATCTTAAACCTTTAAAGTTGCAGAT 2593
 Qy 668 pheProSerThrValThrLeuProLeuAsnArgAsnIleProPheIlePheAnaGala 687
 Db 2594 GGTATACAAACATTAAATTAGCAACAGATAGTCCCTGAGCATGTAACCTAAATTAGGT 2653
 Qy 688 AspValSerAsnSer-----IleLeuIleIleAspLysIleGluPheIlePro 703
 Db 2654 GAAGACCCATTAATTCATATCTGGTATAGTTTACGTTGACCGAATCGAATTCATCCCA 2713
 Qy 704 IleThrSerSerMetHisGlnAenArgGluLysGlnLysLeuGlnThrIleGlnThrLys 723
 Db 2714 GTTAGATGACACATAT-----GAAGCGAACAAGATTAGAACGACGGAAGAAAGCA 2764
 Qy 724 IleAenthRhePheThrAsn 730
 Db 2765 GTGAATGCCTTGTTTACGAAAT 2785

RESULT 12
 US-11-091-643-3
 ; Sequence 3, Application US/11091643
 ; Publication No. US20050246789A1
 ; GENERAL INFORMATION:
 ; APPLICANT: TANAKA, Masao
 ; APPLICANT: YOKOYAMA, Tomoko
 ; APPLICANT: AOYAGI, Moriochi
 ; APPLICANT: HASEGAWA, Makoto
 ; APPLICANT: EHARA, Gaku
 ; APPLICANT: KIMURA, Masaharu
 ; APPLICANT: NISHIHASHI, Hideji
 ; TITLE OF INVENTION: Polypeptide having larvae growth inhibiting or
 ; TITLE OF INVENTION: insecticidal effect on scarabaeidae insects and
 ; FILE REFERENCE: OPI335
 ; CURRENT APPLICATION NUMBER: US/11/091,643
 ; CURRENT FILING DATE: 2005-03-29
 ; PRIOR APPLICATION NUMBER: JP 2001-115754
 ; PRIOR FILING DATE: 2001-04-13
 ; PRIOR APPLICATION NUMBER: JP 2001-203463
 ; PRIOR FILING DATE: 2001-07-04
 ; NUMBER OF SEQ ID NOS: 22
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 3
 ; LENGTH: 4359
 ; TYPE: DNA
 ; ORGANISM: Bacillus popilliae
 ; FEATURES:
 ; NAME/KEY: CDS
 ; LOCATION: (282)..(4229)
 US-11-091-643-3

Alignment Scores:
 Pred. No.: 3,78e-64 Length: 4359
 Percent: 669.00 Matches: 230
 Score Similarity: 43.5% Conservative: 106
 Best Local Similarity: 29.8% Mismatches: 301
 Query Match: 17.3% Indels: 136
 DB: 10 Gaps: 34

US-10-783-417-2 (1-735) x US-11-091-643-3 (1-4359)
 Qy 1 MetAsnGln-----AsnAsnAspAsn-----AenGluTyx 10
 Db 336 ATGATTCATATATCATACCAAAACGATTACAAAAGTTTACACCAAGTGGAAATGAAATG 395
 Qy 11 GluIleIle-----AspSerHisThrSerProTyR-PheProAsnArgAsnSerAsnAsp 28
 Db 396 CAATCATTCACACCTTCAAGTAAAGCTTTACTTACAGTCCCAAT----- 440

Qy 29 SerArgTyProTyR-ThrAsnAsnProAsnGlnProLeuGlnAsnThrAsnTyRysGlu 48
 Db 441 ---AGTATCCGTATCTACGATCCCAATGTCATAGCAGAGGCGGAAAGTTAAATAAT 497
 Qy 49 TrpLeuAsnMetCysGlnGlyAsnThrGlnTyRgLyAspAsnPheGluThrPheAlaSer 68
 Db 498 TGGTGGATATATGTACAGGACA-----GGCGACACAGAAACCCGAAACCTGGT 548
 Qy 69 AlaAspThrIleAlaAlaValSerAlaGlyThrIleValSerGlyThrLeuLeuAlaGly 88
 Db 549 GCTATTTCAAAAGGCTGTCTTCTGCTGCAATTAATTAAGC----- 590
 Qy 89 IleGlyLeuThrSerIleSerGlyProIleGlyIle-----IleGly 103
 Db 591 -----ACCGGGCTTCTGGCTTACAGAGTTCCTGTCATACACAAATCGGG 638
 Qy 104 AlaIleIleIleSerPheGlyThrLeuIleThrValPheTyRProAlaGlyGluAsn 123
 Db 639 GCATTT-----TATACCTTCTTATGAATTAACCTTAAGGCTGCA-----AGC 680
 Qy 124 LysThrValTrpThrGlnPheIleLysMetGlyGluIlePheValAspThrProLeuThr 143
 Db 681 AATTAACAATGGAGAGGAGTTATAGCACATGTGGAAAGAACTCATTAATGCAGAAACCTAAC 740
 Qy 144 GluSerIleLysGlnLeuLysLeuGlnThrLeuGlnGlyPheArgGlnIleLeuGlnSer 163
 Db 741 GATCATGTAGAAATTCGGCACTTACAAATTAATGCTTTACGTAAATACATAGAGATA 800
 Qy 164 TyrAsnThrAlaLeuAspAspTrpArgLysLeuLysArgLeuGlnAlaProGlyLeuPro 183
 Db 801 TATAACGAAGCTTTAATAGTTTGAAACAA-----GATCCT 836
 Qy 184 ProSerSerAlaLeuGlnGlnAlaAlaLeuThrLeuLysIleArgPheGluAsnValHis 203
 Db 837 AACCAATGCAAACTTAAGAT-----GATGTAAGAAATTAATTCGATGGCTCAAT 887
 Qy 204 AsnAspPheIleArgGluIleProGlyPheGlnLeuGlnTyR-TyrIserThrLeuLeu 223
 Db 888 AGCCAAATTCGAAGAAATTAATTCACAAATTTAAAGAAAGGCTTTGAAGTCAATTATTA 947
 Qy 224 ProIleTyRAlaGlnAlaAlaAsnPheHisLeuAsnLeuGlnGlnGlyAlaGluLeu 243
 Db 948 ACTATATATGACCAATCTGCAATCTTCATCTATATTAATTAAGACATTCCTCTTGAT 1007
 Qy 244 AlaAspGluTrpAsnAlaAspIleHisProSerGlnIleGluProAsnAlaGlyThrSer 263
 Db 1008 GGTGCATCTTGGGATTT-----GCTCAAGCTACTATT 1040
 Qy 264 AspAspTyR-TyrLysLeuLeuLysGluAsnIleProLysTySerAsnTyR-CysAlaAsn 283
 Db 1041 GACAAATTAATTCATATCGCAAAATTAAGAAACCGCAGAGTATGCATATCTGTACACT 1100
 Qy 284 ThrTyRArgThrGlyLeuLysAsnLeuArgAspGluProAsnMetLysTrpSerIlePhe 303
 Db 1101 TGGTATCAGACCGGTTTACAAAGATTGCAAGGACACTGCTAGCAGTGGCTCTGTAT 1160
 Qy 304 AsnAspTyRArgArgTyR-MetThrIleThrValLeuAspThrIleSerGlnPheSerLeu 323
 Db 1161 CATGATTTTGAAGAGAAATGACACTTAACAGATTAATGGATTCATTAATTTTCAAT 1220
 Qy 324 TyrAspIleLysArgTyRArgAspSerIleGlyGlyIleGlyValLysGlyIleLysAsn 343
 Db 1221 TATGATGCCGTAGCTTAC-----CCACTGAGGTAAAGGGA----- 1256
 Qy 344 GluLeuThrArgGluIleTyRThrThrGluIleAsn-----PheAspArg 358
 Db 1257 GAGCTTACAGAGGAATTTATACGATCGATGACAGCCGGTACTAATCGATGATGCA 1316
 Qy 359 LeuProGlnLeuArgValGlnProAsnLeuAla---ThrMetGluTyRAsnLeuThrArg 377
 Db 1317 GCACCATCATTCGCAAGAAATAGAAATCTTAATTAATTAAGGCAACCAAGAACTGTACTGG 1376
 Qy 378 AlaSerPheLysLeuPheSerPheLeuGlnGlnPheIlePheTyR-ThrGluAsnThrAsn 397


```

OY 608 AensSerThrgInSerTYrPheHlleArgLeuArgTYrAlaTrpAsnGlyAlaGlyAsnThr 627
Db 1741 ---AAAGCAGGGAATATATCGGTAGACTGAGATATGCTACTGATGCA----- 1785S
OY 628 LeuProAsnIleSerLeuThrIleProGlyValIleGlyIleProProGlnArgLeuAsn 647
Db 1786 -----GATATTGTTGATTGCAATGTAACGATGCTGATTCAGATGCGCAAAAACATGAAAC 1839
OY 648 AsnThrPheSerGlyThrAsnTYrAsnAsnLeuGlnTYrGlyAspPheGlyTYrPheGln 667
Db 1840 CCAGGT-----GAGGATCTGCATCATTAACCTTTAAAGTTGCAGAT 1881
OY 668 PheProSerThrValThrLeuProLeuAsnArgAsnIleProPheIlePheAsnArgAla 687
Db 1882 GCTATCAACAACAGTAAATTATGACCAACAGATGTCGGTGGACGATTAAACATTAATTAGGT 1941
OY 688 AspValSerAsnSer-----IleLeuIleIleAspValIleGluPheIlePro 703
Db 1942 GAAGACCTTAATTCAACATTTATCTGTTGTTATGTTACGTTGACCAATGATTCATCCCA 2001
OY 704 Ile 704
Db 2002 GTA 2004

RESULT 15
US-11-058-727-79
; Sequence 79, Application US/11058727
; Publication No. US20050261483A1
; GENERAL INFORMATION:
; APPLICANT: Andre R. Abad
; APPLICANT: Ronald D. Flannagan
; APPLICANT: Rafael Herrmann
; APPLICANT: Theodore W. Kahn
; APPLICANT: Albert L. Lu
; APPLICANT: Billy Fred McCutchen
; APPLICANT: James K. Premeau
; APPLICANT: James F.H. Wong
; APPLICANT: Cao-Guo Yu
; TITLE OF INVENTION: Genes Encoding Proteins With Peestcidal
; TITLE OF INVENTION: Activity
; FILE REFERENCE: 35718/287809
; CURRENT APPLICATION NUMBER: US/11/058,727
; PRIOR FILING DATE: 2005-02-15
; PRIOR APPLICATION NUMBER: 60/391,786
; PRIOR FILING DATE: 2002-06-26
; PRIOR APPLICATION NUMBER: 60/460,787
; PRIOR FILING DATE: 2003-04-04
; PRIOR APPLICATION NUMBER: 10/606,320
; PRIOR FILING DATE: 2003-06-25
; NUMBER OF SEQ ID NOS: 134
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 79
; LENGTH: 2025
; TYPE: DNA
; ORGANISM: Bacillus thuringiensis (mutated)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)....(2025)
US-11-058-727-79

Alignment Scores:
Pred. No.: 2,15e-63 Length: 2025
Score: 657.50 Matches: 212
Percent Similarity: 44.88 Conservative: 120
Best Local Similarity: 28.64 Mismatches: 292
Query Match: 17.08 Indels: 117
DB: 8 Gaps: 31

US-10-783-417-2 (1-735) * US-11-058-727-79 (1-2025)
OY 5 AsnAspAsnAsnGluTYrGluIleIleAspSerHisThrSerProTYrPheProAsnArg 24
Db 10 AATATATCAAAATGATATGAATATATATGATCGACACCTCTT-----ACTTCT 57

```

OY	25	AmsberlsnApsSer---	ArgTyrProDTrThrsAmnProAnGlnProLeuGlnAsn	43
Db	58	GTATCCATGATGTTCTAACAGATACCTTTTTCGATATAGCCACAAATGCGCTACAAAT	117	
OY	44	ThrsAntYrLysGluTrpLeuAsnMetCysEngInYAsn---	ThrsInYrGLYAspAsn	62
Db	118	ATGATATTATAAGATTATTATAAAATGTCTCGGGAAATGTAAGTAATCCCTGGTCA	177	
OY	63	PheGluThrPheAlaSerAlaAspThrIleAlaIaValSerAlaGlyThrIleValSer	82	
Db	178	CCTGAAGTACTGTTAGCGGACAAAGAT-----GCAGCTAAGCGCGCAATGATATAGTA	231	
OY	83	GlyThrLeuLeuAlaGlyIleGlyGlyLeuThrserIleSerGlyProIleGlyIleIle	102	
Db	232	GGTAAATTACTATCAGGTTTAGGG-----GTTCCCATTTGTT	267	
OY	103	GlyAlaIleIleIleSerPheGlyThrLeuIleThrValPheTrpProAlaGlyGln	122	
Db	268	GGCGCGATAGTGAAGCTTTTATCTCAACTTATATGATTCTGTGGCCTTCAGGGGAA---	324	
OY	123	AspYrThrValTrpThrGlnPheIleLysMetGlyGluIlePheValAspThrProLeu	142	
Db	325	---AAGAGTCAATGGGAAATTTTATGGAACAATGAGAAGACATTAATCAAAAAATA	381	
OY	143	ThrGluSerIleLysGlnLeuLysLeuGlnThrLeuGluGlyPheArgGlnIleLeuGln	162	
Db	382	GCAGAAATATGCAAGAAATAAAGCCCTTTCCGAAATTAGAGATTAAGTAATTAATCAAA	441	
OY	163	SerYrThrAntPrAlaLeuAspAspTrpArgLysLeuLysArgLeuGlnAlaProGlyLeu	182	
Db	442	TTTATATCTAATGTGGCTTGGAAGATGGAGAA---AATCCATTTCCAAAGTCAGGTCCA	498	
OY	183	ProProSerSerAlaLeuGlnGlnAlaAlaLeuThrLeuLysIleArgPheGlnAsnVal	202	
Db	499	AATGATCCCGGGCGCTTACAGAT-----GTCCGAATCCATTTGAATCTGTG	546	
OY	203	HisAsnAspPheIleArgGluIleProGlyPheGlnLeuGluThrYrLysThrLeuLeu	222	
Db	547	GATAGTTATTAGCAATATATATGCCATCTTTTAGAGTGAACAATTTGAAGTACCAATTC	606	
OY	223	LeuProIleYrAlaGlnAlaIaAsnPheHisLeuAsnLeuLeuGlnGlnGlyAlaGlu	242	
Db	607	CTTACTGATATGCAATGCGAGCCACCACTTACTTGTATTAAGACGCGCTCAATT	666	
OY	243	LeuAlaAspGluTrpAlaAlaAspIleHisProSerGlnIleGluProAlaAlaGlyThr	262	
Db	667	TTTGGAGAGAGATGGGAGATG-----TCAACAATCTACT	699	
OY	263	SerAspAspYrYrYrLysLeuLeuLysGluAsnIleProLysYrSerAsnYrCysAla	282	
Db	700	ATTATATACTATTATGATTCGTCAAATGAACCTTACTGCAAAATTTCTGATCACTGTGA	759	
OY	283	AsnThrYrArgThrGlyLeuLysAsnLeuArgAspGluProAsnMetLysTrpSerIle	302	
Db	760	AAGTGTATGAAATCGGTTTAGCAAAATTAAGACGACGACGCTAAACATGCGTTGAC	819	
OY	303	PheAsnAspYrYrArgArgYrMetThrIleThrValLeuAspThrIleSerGlnPheSer	322	
Db	820	TATATCAATTCCTGTAGGAATATACACTGGCGGTTTTAGATGTTGTGTGCTATTATCCCA	879	
OY	323	LeuYrAspIleLysArgYrYrArgAspSerIleGlyGlyIleGluValLysGlyLys	342	
Db	880	AATTATGACACAGCAGGTAC-----CCATGGAAGAAAGCA-----	918	
OY	343	AsnGluLeuThrArgGluIleYrYrThrThrGluIleAsnPheAspArgLeuProGlnLeu	362	
Db	919	---CAACTAACAGGAAGTATATACA-----GATCCACTGGCGGCGCGTA	960	
OY	363	ArgValGln-----ProAsnLeuAlaThrMetGluYr	373	
Db	961	AACGTGTCTTCATTTGGTTCCTGTATGACAAAGCACTTTTCGAGATGATGAATCA	1020	

QY 374 AsnLeuThrArgAlaSerPheLeuPheSerPheLeuGlnPheIlePheTyrThr 393
Db 1021 TCCGTTATTCCGACCAACC---CATGATTTGATTATATAACGGACCTCAGCTGATATCA 1077
QY 394 GluAsnThrAsnPheGlyAsn---ArgLeuVal-----GlyIleSer 406
Db 1078 CATCAAGAGATTCCTTCGCTGATATATAAGCATTTGGCTGTCATCAATAAGC 1137
QY 407 AsnArgAspAlaProThrTyrSerAsnThrIleThrGluThrLeuTyrGlyIuArgThr 426
Db 1138 TACCATCGGTGACAGTAGGGGTAGTAT-----CTTCAACAATGATGAACTATCA 1191
QY 427 GlySerProThrThrIleThrThrIleArgProPheGluSerTyrIleValSerIleValThr 446
Db 1192 AATCTACACAGCACTAGTACTTT---GATTTACCAATATGATATTTACAGACTCTA 1248
QY 447 AspArgGlnSerProProValSerProIleGlnPro-----HisPheIle 461
Db 1249 TCAAAGATGACGATCTCTGATATATTGTTTACCTGGTATACGTATATATTTTGGGA 1308
QY 462 IleAsnGlnIleGluLeuThrLeuAsnGlySerSerAsnAsn-----ThrLeuIys 478
Db 1309 ATGCCAAGAGTCAAGTTTTCATGTAACCAATTGAATATACCAAGAAAGCGTTAAAG 1368
QY 479 TyrSerAlaGlyGlySerLeuSerAsnTyrGlnAsnThrThrPhePheGlnPheProArg 498
Db 1369 TAT-----AATCCAGTTTCCAAGATATTTATAGCAGTACAAGA 1407
QY 499 IysIysAspCysAsnLeuValIleAspProGlyCysSerProAsnPheAsnAsnTyrSer 518
Db 1408 GATTCGGAATTAGATTAATCTCCTGATTAATCTGATTAATCTGATTAATCTGATTAATC 1467
QY 519 HisIleLeuSerHisPhe---SerLeuPheThrTyrSerTyrValIleGlyLeuGlnLeu 537
Db 1468 CATAGATTATGTCATATCAAGTATCCCGGACCGGTAAACATACCGGATTA----- 1521
QY 538 GlnIleLeuAspThrGlyValLeuGlyTyrPheHisSerSerValAspArgTyrAsnAla 557
Db 1522 -----GTACCTGATTTTCTTGACACATCGAAGTGCAGATTTTAAACAATACA 1569
QY 558 IleSerAspIysIleIleThrMetIleProAlaIleIys---GlyAsnAsnLeuAspThr 576
Db 1570 ATATATTCAGATAAATCACTCAAAATTCGCGCGTTAAATGTTGGATTAATTTACCGTTT 1629
QY 577 AsnSerIysValIleGluGlyProGlyHisThrGlyGlyAsnLeuVal---TyrLeuGln 595
Db 1630 ---GTTCCAGTGTAAAGACCAAGACATACAGAGGGGATTTATACAGTATATAGA 1686
QY 596 SerGlnGly-----ArgLeuGlnIleThrCysGluThrPro 607
Db 1687 AGTACTGCTTGTAGGAACCTTATTTTACGCTCGATATGCGCTAGCATTAAGAA----- 1740
QY 608 AsnSerThrGlnSerTyrPheIleArgLeuArgTyrAlaThrAsnGlyValGlyAsnThr 627
Db 1741 ---AAAGCAGGAAATATCGTAAAGACTGAGATATGCTACTGATGCA----- 1785
QY 628 LeuProAsnIleSerLeuThrIleProGlyValIleGlyIleProProGlnArgLeuAsn 647
Db 1786 -----GATATTGTATTGATGATAAACGATGCTCAGATTCAGATGCCAAACATGAAAC 1839
QY 648 AsnThrPheSerGlyThrAsnTyrAsnAsnLeuGlnTyrGlyAspPheGlyTyrPheGln 667
Db 1840 CCAAGT-----GAGGATCTGACATCTAAACCTTAAAGTTGCAGAT 1881
QY 668 PheProSerThrValThrLeuProLeuAsnArgAsnIleProPheIlePheAsnArgAla 687
Db 1882 GCTATCACAAAGATTATTAATTAACAACAGTAGTTCGTTGCAGTTAAACATATTTAGGT 1941
QY 688 AspValSerAsnSer-----IleLeuIleIleAspIysIleGluPheIlePro 703
Db 1942 GAAGACCTTAATTAATTAATTAATGATTAATGATTAATGATTAATGATTAATCAATCA 2001
QY 704 Ile 704

Db 2002 GTA 2004

Search completed: January 20, 2006, 23:03:36
Job time : 382 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - nucleic search, using frame_plus.p2n model

Run on: January 20, 2006, 17:27:41 ; Search time 1439 Seconds

(without alignments)
4223.762 Million cell updates/sec

Title: US-10-783-417-2

Perfect score: 3869

Sequence: 1 MNQNNNDNNVEYLIDSHSTPY.....KLETIQTQKINPTNHTKTL 735

Scoring table:
BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 9793542 seqs, 4134689005 residues

Total number of hits satisfying chosen parameters: 19587084

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODE=frame+g2n.model -DEV=x1h
-O=/cgn2_1/USPTO_SPOOL/US10783417/runat_20012006_095011_23921/app_query.fasta_1.903
-DB=Published Applications NA Main -QFMT=fastap -SUFFIX=p2n.rnpbm
-MINMATCH=0.1 -LOOPEL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1
-MATRIX=Dlosum62 -TRANS=human40.cdi -LIST=45 -DOCCALIGN=200 -THR_SCORE=pct
-THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORR=ext
-HEADSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US10783417@cgn_1_1.1026 @runat_20012006_095011_23921 -NCPU=6 -ICPU=3
-NO MMAP -NEG SCORES=0 -WAIT -DSBLOCK=100 -DEV TIMEOUT=120 -WARN TIMEOUT=30
-THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7 -YGAPOP=10
-YGAEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications NA Main:*

1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
3: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:*
4: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*
5: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
6: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
7: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
8: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq:*
9: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq:*
10: /cgn2_6/ptodata/1/pubpna/US11_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3869	100.0	2208	8	US-10-783-417-1
2	3314	85.7	2235	8	US-10-782-570-1
3	3104.5	80.2	2085	8	US-10-782-570-3
4	849	21.9	4896	3	US-09-756-526A-3
5	849	21.9	4896	6	US-10-345-020-3
6	849	21.9	4896	6	US-10-342-821-3
7	759	19.6	3684	9	US-10-929-754-2

8	705	18.2	4391	9	US-10-929-754-4	Sequence 4, Appl1
9	703	18.2	3633	5	US-10-032-717-3	Sequence 3, Appl1
10	703	18.2	3633	6	US-10-414-637-3	Sequence 3, Appl1
11	703	18.2	3633	7	US-10-606-320-3	Sequence 3, Appl1
12	703	18.2	3633	8	US-10-746-914-3	Sequence 3, Appl1
13	703	18.2	6613	5	US-10-032-717-28	Sequence 28, Appl1
14	703	18.2	6613	6	US-10-414-637-28	Sequence 28, Appl1
15	703	18.2	6613	6	US-10-606-320-18	Sequence 18, Appl1
16	703	18.2	6613	8	US-10-746-914-18	Sequence 18, Appl1
17	689.5	17.8	4188	5	US-10-120-544A-5	Sequence 5, Appl1
18	689.5	17.8	4188	10	US-11-091-654-5	Sequence 5, Appl1
19	678.5	17.5	2003	5	US-10-032-717-7	Sequence 7, Appl1
20	678.5	17.5	2003	6	US-10-414-637-7	Sequence 7, Appl1
21	678.5	17.5	2022	5	US-10-032-717-17	Sequence 17, Appl1
22	678.5	17.5	2022	6	US-10-414-637-17	Sequence 17, Appl1
23	678.5	17.5	2022	7	US-10-606-320-13	Sequence 13, Appl1
24	678.5	17.5	2022	8	US-10-746-914-13	Sequence 13, Appl1
25	675	17.4	3621	5	US-10-032-717-1	Sequence 1, Appl1
26	675	17.4	3621	6	US-10-414-637-1	Sequence 1, Appl1
27	675	17.4	3621	7	US-10-606-320-1	Sequence 1, Appl1
28	675	17.4	3621	8	US-10-746-914-1	Sequence 1, Appl1
29	675	17.4	3621	10	US-11-021-115-5	Sequence 5, Appl1
30	675	17.4	4874	5	US-10-032-717-27	Sequence 27, Appl1
31	675	17.4	4874	6	US-10-414-637-27	Sequence 27, Appl1
32	675	17.4	4874	7	US-10-606-320-17	Sequence 17, Appl1
33	675	17.4	4874	8	US-10-746-914-17	Sequence 17, Appl1
34	669	17.3	4359	5	US-10-120-544A-3	Sequence 3, Appl1
35	669	17.3	4359	10	US-11-091-654-3	Sequence 3, Appl1
36	668	17.3	2088	10	US-11-018-615-5	Sequence 5, Appl1
37	663	17.1	1959	7	US-10-614-076-37	Sequence 37, Appl1
38	663	17.1	2358	10	US-11-018-615-28	Sequence 28, Appl1
39	663	17.1	3510	10	US-11-018-615-24	Sequence 24, Appl1
40	660	17.1	1959	5	US-10-232-665-3	Sequence 3, Appl1
41	658	17.0	1959	7	US-10-614-076-27	Sequence 27, Appl1
42	658	17.0	1959	7	US-10-614-076-35	Sequence 35, Appl1
43	657.5	17.0	2025	7	US-10-606-320-73	Sequence 73, Appl1
44	657.5	17.0	2025	7	US-10-606-320-79	Sequence 79, Appl1
45	657.5	17.0	2025	8	US-10-746-914-73	Sequence 73, Appl1

ALIGNMENTS

RESULT 1
US-10-783-417-1
; Sequence 1, Application US/10783417
; Publication No. US20040216186A1
; GENERAL INFORMATION:
; APPLICANT: Carozzi, Nadine
; APPLICANT: Harzies, Tracy
; APPLICANT: Kozziel, Michael G.
; APPLICANT: Duck, Brian
; APPLICANT: Carr, Brian
; TITLE OF INVENTION: AXMI-006, A Delta-Endotoxin Gene and
; FILE REFERENCE: 045600/274146
; CURRENT APPLICATION NUMBER: US/10/783,417
; CURRENT FILING DATE: 2004-02-20
; PRIOR APPLICATION NUMBER: 60/448,806
; PRIOR FILING DATE: 2003-02-20
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 2208
; TYPE: DNA
; ORGANISM: Bacillus thuringiensis
; FEATURES:
; NAME/KEY: CDS
; LOCATION: (1)...(2208)
US-10-783-417-1
Alignment Scores: 0
Pred. No.: 3869.00
Length: 2208
Matches: 735

Percent Similarity: 100.0%
Best Local Similarity: 100.0%
Query Match: 8
Gaps: 0

US-10-783-417-2 (1-735) x US-10-783-417-1 (1-2208)

1 MetAaNgInAaenAaSpaNaNgInUTrgJuIlleIaSpSeRHeIeThSeRProTyR 20
1 ATGAATCAAAATTAACGATTAATACGAATATGAAATTAATGATTCGATACCTCAT 60
21 PhePAAaNgArGaSeRSeRaSpSeRArgTyRProTyRThAaAaSPProaNgInPro 40
61 TTTCCGAACAGAAACGTAATGATTCGATACCTTCACAAATTAATCCAAATCAACA 120
41 LeuGInAaNTHeRSeNTyRleSeGInUTrPLeuAaNeMcySGInGlyAaSeNTHeRgInTyRgLy 60
121 TTAACAACACAAATTAACAAGAGTGGCTCAATATGTCAGAGGAATACAAATATGCT 180
61 AaPaenPheGInUThrPheAlaSeRaIaSPThRIaIaIaValaSeRaIaGlyThRIle 80
181 GATAATTCGAGACATTGCTAGTCTGATACAAATGGCGAGTTAGTCAAGTACTAT 240
81 ValSeRgLyThrLeuLeuAlaGlyIleGlyGlyLeuThrSeRIleSeRgLyProIleGly 100
241 GATCCGGTACTCTGTAGCCGGTATAGTGGGCTCATTCTATATCCGGAACCGATAGA 300
101 IleIleGlyAlaIleIleSeRPhEGlyThrLeuIleThValaPheTriProAlaGly 120
301 ATTAATAGGTGCTATTAATATATCTTTGGTACCTTAATCACTGCTTTGGCCCGGGA 360
121 GluGInAaPlyeThrValITrPThrGInPheIleIySmeGlyGluIlePheValaSPThr 140
361 GAACAAGACAAACAGTATGACACAAATTTAATAAGGAGAAATTTTGTCATACA 420
141 ProLeuThrGInUSeRIleIySeGInleuIySeLeuGInThrLeuGInGlyPheArgGInIle 160
421 CCGTTAACAGAAAGCATAAACAGCTAAAGTTAACAACTTTAAGAGATTAAGACAAATA 480
161 LeuGInSeTyRSeNTHeRAlaLeuAaSPeRTpARgLySeLeuIyARgLeuGInAlaPro 180
481 TTAACAAGCTATTAATACACATTAAGTATGGAAGAAATTAATAAGACTACACACTGCT 540
181 GlyLeuProProSeSeRaIaLeuGInGInAlaIaLeuThrLeuIyIleArgPheIle 200
541 GGATTAACCAACCATCAACGACATTACAAACAGCTGCTGACCTTAATAATAGATTAG 600
201 AaNTValIHaenAaSPheIleArgGluIleProGlyPheGInleuGInUThrTyRlySeThr 220
601 AATGTTCACAATGATTTATTCGAGAAATATCCTGGTTCCAACTTGAAACTTATATAACG 660
221 LeuLeuLeuProIleTyRAlaGInAlaIaAaenPheIleAaenLeuLeuGInGInIly 240
661 CTATTACTACTTATTAATGCGAGCTGCTAATTTTCATTTAATTTATTAACAACAGCT 720
241 AlaGInleuAlaAaSPGInUTrPaSnaIaSPleIleIaSProSeRgInIleGInUProaNaIa 260
721 GGTGAATGGCTGATGAGATGAGATCAAGATATACATCCCTCAAAATGAACCTAATGCT 780
261 GlyThrSeRaSPeRTyRlyIySeLeuLeuIySeGInAaenIleProIySeTyRSeRaNTyR 280
781 GGAATCATGATGATCATTTATTAACCTTTAAAGAAATATATACCTTAATATATGTAACAT 840
281 CysAlaAaenThTyRArgThrGlyLeuIyAaenLeuAaSPGInUProaNeMcySeThr 300
841 TGTCGAATTAATCTTAATAGAACAGAGCTAAATAATCTTAAGAGCAACAAATATGAATGG 900
301 SerIlePheAaenAaSPTyRArgTyRSeThRIleThValaIleuAaSPThRIleSeRgIn 320
901 AGTATATTTAATGATGATGAGATATATGACCATTAATGATTAATCACTCTCTCA 960
321 PheSeRleuTyRAsPIleIyAaRgTyRArgAaSPSeRIleGlyIleIleGluValIySeGly 340

961 TTTCTTATATATGATATATAAAGATATAGACATTCATAGAGCAATAGAAATAAGCC 1020
341 IleIySaNGInleuThrArgGluIleTyRThrThrGluIleAaSPheAaSPArgLeuPro 360
1021 ATTAAGATGAACCTCACAGAGAAATTTATACACTGAAGTAATTAATTTGATCGTCTTCT 1080
361 GInleuArgValGInProAaenLeuAlaThMeGluTyRAsnleuThrArgAlaSeRPh 380
1081 CAACTAGAGTTCACCAACCAATCTAGCTAGATGGAATTAATTTAACACAGTCAAGTTT 1140
381 LysleuPheSeRPhleuGInPheIlePheTyRThrGInAaenThPhePheGlyAaen 400
1141 AAATTAATTTATATTTTAAGCAATTAATTTTATATACAGAAATTAACAAATTTCCGGAAT 1200
401 ArgLeuValGlyIleSeRaNaRgAPAlaProThTyRSeRaenThRIleThrGInUThr 420
1201 CGTTAGTGGTATTTCTAATCGATGACACTTATATACCATCTATACATGAACT 1260
421 LeuTyRgLyGInUArgThrGlySeRProThrThrIySeThRIleArgProPheGInSeTyR 440
1261 TTATATGAGAGAAAGACAGTTTCAACCAACAAACAAATTAAGACATTGGAATCTTAT 1320
441 LysValSeRIleValIThrAsPaRgGInSeRProProValSeRProIleGInProHIsPhe 460
1321 AAGTTTCAATTTGAATGATGATAGACAAATCACCTCTGTTCCCTTAATTCACACACTTT 1380
461 IleIleAaNGInIleGInleuTyRleuAaNGIlySeSeRaAaenThrLeuIySeTySeR 480
1381 ATTAATTAATCAAAATTTGAACCTTATTTAATATGGCTATTAACACACACTCAAAATTTCA 1440
481 AlaGlySeRleuSeRaNTyRGIaSeNThrPhePheGInPheProArgIyLeIyS 500
1441 GCAGAGGGCTTTATATCAATTAATCAAAACAACTTTTTCATTTCTTAAGAAAAA 1500
501 AaPcySaenleuValIleAaSProGlycySeSeRProAaPheAaenTySeThRIle 520
1501 GACTGCAATCTAGATTAATGATTCAGTGGTTCACCAAACTTTAATTAATGATATAT 1560
521 LeuSeRHisPheSeRleuPheThrTySeRtyValIleGlyLeuGInleuGInIleu 540
1561 TTATCCCATTTTTCATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1620
541 AaPThrGlyValleuGlyTyRThRIleSeSeRaValaAaPaRgTyRAsnaIleSeSeRaP 560
1621 GATACAGGTGATTAAGATGACACACAGTGGTGTGATATTAATTAATCAATATCAAT 1680
561 LysIleIleThrMeIleProAlaIleIySeIyAaenAaenleuAaSPThRAsnSeTySeVal 580
1681 AAAATTAATTAATGATTCACAGCAATCAAAAGTAACAATCTTGATTAACAACCTTAAGGTA 1740
581 IleGInUlyProGlyHIsThrGlyIyAaenleuValTyRleuGInSeRgInGlyArgLeu 600
1741 ATTGAAGACCTGGTCAATACAGAGAAACTTGGTTTATTTTAACAAGTCAAGGGCTTTA 1800
601 GluIleThrcySGInUThrProAaenSeThrGInSeTyRPhelIeArgLeuArgTyRAla 620
1801 GAAATTAATGATGAACCTCAATTCACACAACTTATTCATTAAGACTCGATATGCT 1860
621 ThrAaNGlyAlaGlyAaenThrleuProAaenIleSeSeRleuThRIleProGlyValIleGly 640
1861 ACAAATGAGTGTGAAATATCTTCCTAATATATCTTCAATTAACAGAGGATTAATAGA 1920
641 IleProProGInArgLeuAaAaenThrPheSeRgLyThrAaNTyRAsnaSPleuGInTyR 660
1921 ATACCACTCAACGACTCAACACACTTTTCTGCTGATCAAAATTAATTAATTAATTAAC 1980
661 GlyAaSPheGlyTyRPhEGInPheProSeRThrValThrleuProleuAaNaRgAaIle 680
1981 GGAATTTTGGGTATTTCCAAATTTCCAAAGTACAGTAACATTAACCTTTAATCGAAACATA 2040
681 ProPheIlePheAaNaRgAlaAaPyaISeRSeNTyRleuIleIleAaPlyIleGlu 700
2041 CCATTAATTAATTAATCGGAGATGATATCAAAATTTAATCAATTAATTAATTAATTAAT 2100

Db	358	GAACAGACAAACAGATGACACATTTATTAATGGAGAAATTTTGTGTGATACA	417
OY	141	ProLeuThrgIuSerIleIyegInLeuIySeuInThrLeuGluIyPheArgGlnIle	160
Db	418	CCGTTAACAGAAACATATAAACAGCTTAAGTTACAACTTTAGAAGATTAGACAATA	477
OY	161	LeuGlnIeSerTyrrAntThrAlaLeuApsrTPArgIySeuIySaGleuGlnAlaPro	180
Db	478	TTACAAAGCATAATATACAGCATTAAGATGGAGAAATTTAAACATACAGAGCTCT	537
OY	181	GlyLeuProProSerSerAlaLeuGlnAlaIaLeuThrIleuIySileArgPheGlu	200
Db	538	GGATTACACCATCATACGACATTCACAAAGCTCCTTGACTCTTAATAATACGATTTGAG	597
OY	201	AsnValHisAsnApsPheIleArgIuIleProGlyPheGlnLeuGluThrTyIyLeuThr	220
Db	598	AATGTCACATATGATTTTATTCGAGAAATACCTGTTCCAACTTGAACTTATTAACAG	657
OY	221	LeuLeuLeuProIleTyrrAlaGlnAlaIaAsnPheHisIleuAsnLeuLeuGlnGlnIy	240
Db	658	CTATTACTACTATTTATTCGCAAGCTGCTAAATTTTCATTTAAATTTATACAACAAGT	717
OY	241	AlaGlnLeuAlaAspGluTPAsnAlaAspIleHisProSerGlnIleGluProAsnAla	260
Db	718	GCTGAATTCGCTGATGCAATGCAATGCAATATACATCTTCACAAATGGAACCTTAAGCT	777
OY	261	GlyThrSerApsApsTyrrTyIyLeuLeuIySeuGlnAsnIleProIyTySerAsnTyrr	280
Db	778	GGAACATCAGATGACTTATTAACCTTTTAAAGAAATATACCTAATATATAGTACATAT	837
OY	281	CysAlaAsnThrTyrrArgThrGlyIleuIyAsnLeuArgAspGluProAsnMetIyTrp	300
Db	838	TGTCCAATATCTTATACAGAAAGCATTAATTAACCTTGAAACGAACTTAATATAGATGAG	897
OY	301	SerIlePheAsnApsTyrrArgArgTyrrMetThrIleThrValIeAspThrIleSerGln	320
Db	898	AGTATATTTAATGATATTCAGAGATATATATGCTATGATATATGATACATCTGCTCA	957
OY	321	PheSerLeuTyrrApsIleIyAsnTyrrArgApsSerIleGlyIleGluValIyGly	340
Db	958	TTTTCTTTATATGATATTAAGAGATACAAAGATTCAATAGCA-----AGAAATAGTGGC	1011
OY	341	IleIyAsnGluLeuThrArgGluIleTyrrThrThrGlnIleAsnPheApsArgLeuPro	360
Db	1012	ATTAATACTGAACCTTACAGAGAAATTTATACAACTGAATTAATTTTGACCGCTTACT	1071
OY	361	GlnLeuArgValGlnProAsnLeuAlaIaThreGlnTyrrAsnLeuThrArgAlaSerPhe	380
Db	1072	TACCTTGAAATCTCAACCAATCTCGCATATATGAAATATATTAACACGTCAGGAGCTT	1131
OY	381	IySeuPheSerPheLeuGlnGlnPheIlePheTyrrThrGlnAsnThrAsnPheGlyAsn	400
Db	1132	AGATTATTTTCATTTTATAGATGAACTTATATTTTATCAAAAATGAAAGTACGGGAAT	1191
OY	401	ArgLeuValGlyIleSerAsnArgApsAlaProThrTyrrSerAsnThrIleThrGluThr	420
Db	1192	CGTTAGTGGTATTTGGAATCGTAATAGATATCTATATGCTACAGACAGAACTGAATAT	1251
OY	421	LeuTyrrGlyIuAqThrGlySerProThrThrIySerThrIleArgProPheGlnSerTyrr	440
Db	1252	ATTAATATGAGAAACAGGTCACCCACAAACAAACCTTAATATACATTTGAATCTCAT	1311
OY	441	IyValSerIleValThrApsArgGlnSerProProValSerProIle---GlnProHis	459
Db	1312	AAATTTTCATTTGTAATGATAGACATGAACTCTACCTCCCTTTCTACATATATAC	1371
OY	460	PheIleIleAsnGlnIleGluLeuTyrrLeuAsnGlySerSerAsnAsnThrLeuIyTyrr	479
Db	1372	TTTACATTAATACAAATTTGAACCTTATTTAAATTAATCACCTAGTATTAATTAACATAT	1431
OY	480	SerIleIleGlySerLeuSerAsnTyrrGlnAsnThrThrPhePheGlnPheProArgIyS	499
Db	1432	TCAGCTGGGGGGAATTTATCTAAATGATATAAAAAACAATGATTTTCAATTTCTGTAAAA	1491


```

QY 500 LysaapCyasentleuvalilleapProgllyCySerProasnpheasnsntYzSerHis 519
Db 1492 AAAGACTGTAACCAATTATTAATCAAAATGTTTACCAAGCTATTAATGTTATAGTCAT 1551
QY 520 ILeuSerHisPheSerleupheThrTySerTyValIleGlyLeuGlnLeuGlnIle 539
Db 1552 ATTATATCCAGATTTCTTATTTATTAATTAATTCCTTAATAAATGGATTAAGCCGTAATATA 1611
QY 540 LeuAepThrGlyValIleuGlyTyThrHisSerSerValAspArgTyAsnAlaIleSer 559
Db 1612 TTATATACAGGTGATTCAGATGAGACACACACTAGTTTAATAGAAATATATCAATATCA 1671
QY 560 AsplysIleIleThrMetIleProAlaIleYsgIyaasnsleuaspThrasnserys 579
Db 1672 GATAAATAATATACATGATCCACAGCAATCAAAAGTAACAGCTTGATTAACAACCTAAG 1731
QY 580 ValIleGlyGlyProGlyHisThrGlyGlyAsnleuValTyThrLeuGlnSerGlnIlyArg 599
Db 1732 GTAATTGAAGACCTGGTCAATACAGAGGAACTTGTTATTTTCAAAAGTCMAAGGCGCT 1791
QY 600 LeuGluIleThrCyGluTyThrProasnserThrgInserTyPheIleArgLeuArgTy 619
Db 1792 TTAGAGATTACATGATGAACCTCTTAATCTTACACAATCTTATTAATTAAGCTTCAATTC 1851
QY 620 AlaThrAsnGlyAlaGlyAsnThrLeuProAsnIleSerleuThrIleProGlyValIle 639
Db 1852 GCTACAAATGGTGCGGAAATACCTCTTCTAATATATCTTACAAATACAGAGATATAA 1911
QY 640 GlyIleProProGlnIleArgLeuAsnAsnThrPheSerGlyThrAsnTyAsnAsnLeuGln 659
Db 1912 GGAATACCACTGCAACGACTCAACAACCTTTTCTGGTACAAATTAATTAATTAATTAACA 1971
QY 660 TyrgIyaapPheGlyTyThrPheGlnPheProserThrValThrLeuProleuAsnArgAsn 679
Db 1972 TACGAGATTTTGGTATTTCCAAATTCAGATACAGTAACCTTACCTTTAATGGAATC 2031
QY 680 ILeuProheIlePheAsnArgAlaAspValSerAsnSerIleLeuIleIleAspIlySile 699
Db 2032 ATACCATTTATATTAATTCGTCGACATGATATCAATTCATTTATATCATGATTAATTT 2091
QY 700 GluPheIleProIleThrSerSerMetHiseGlnAsnArgIlyGlnIlyLeuGlnIlyThr 719
Db 2092 GAATTTAATCAAAATTAATCTCTCTGTGCGCCAAATTAAGAAAAAATTAAGAAACT 2151
QY 720 ILeuGlnThrIlyIleAsnThrPhePheThrAsnHisThrIlyS 723
Db 2152 ATCCAAACAAAATAATATCAATTTTTCACAAATCATACAAA 2193

```

```

; NAME/KEY: CDS
; LOCATION: (1) ... (2085)
; US-10-782-570-3

Alignment Scores:
Pred. No.: 0 Length: 2085
Score: 3104.50 Matches: 598
Percent Similarity: 91.94 Conservative: 30
Best Local Similarity: 87.64 Mismatches: 52
Query Match: 80.24 Indels: 3
DB: 8 Gaps: 2

US-10-783-417-2 (1-735) x US-10-782-570-3 (1-2085)
QY 52 MetCyGlnGlyAsnThrGlnTyrcIlyAspAsnPhelGlnThrPheAlaSerAlaAspThr 71
Db 1 ATGTGTCAAGGGAATACCAATATGTGTATATTTTCAGACACTTGGCTAGTGTATACA 60
QY 72 ILeuAlaValSerAlaGlyThrIleValSerGlyThrLeuLeuAlaGlyIleGlyGly 91
Db 61 ATTCGTGAGTTAGTGACAGTACTATGTATCCGGTACTGTGTAGCCGTTATAGTGGG 120
QY 92 LeuThrSerIleSerGlyProIleGlyIleIleGlyValaIleIleIleSerPheGlyThr 111
Db 121 CTCACCTTATATCCGACCGATAGGAATATAGGTGTATTAATTAATCTTTGGTACC 180
QY 112 LeuIleThrValPheTrpProAlaGlyGlyGlnIleAspIlySthrValTrpThrGlnPheIle 131
Db 181 CTATACACTGCTCTTTGGCCCCCGGAGAGACACAGCAACAGTATGACACAAATTTAT 240
QY 132 LysMetGlyGluIlePheValAspThrProleuThrgInserIleYsgIleuIlyLeu 151
Db 241 AAAATGGGAGAAATTTTGTGTATACCCGTTAACAGAAACATATAACAGCTTAAGATTGG 300
QY 152 GlnThrLeuGlyGlyPheArgGlnIleLeuGlnIleTyraAsnThrAlaLeuAspAspTrp 171
Db 301 CAACCTTGAAGAGATTGAGCAAAATATTACAAAGCTTAATTAATGATGATTGG 360
QY 172 ArgIlySleuIlyAspGlyLeuGlnAlaProGlyIleuProProSerSerAlaLeuGlnIlyAla 191
Db 361 AGAATAATTAAAGACTACCAAGCTCTCGATTTACACCATCATCGATTTACACAACT 420
QY 192 AlaLeuThrLeuIlyIleArgPheGlnAsnValHisAsnAspPheIleArgGluIlePro 211
Db 421 GCCTTGACTCTTAAATTCGATTTGAGAAATGTTCCAAATGATTTTATTCGAGAAATFACCT 480
QY 212 GlyPheGlnLeuGlnIlyThrTyThrLeuLeuIleuProIleTyAlaGlnAlaIleAsn 231
Db 481 GGTTCCAACTTGAAACTTATTAACGCTATTTACTTATTTATGCGCAAGCTGTAT 540
QY 232 PheHisLeuAsnLeuLeuGlnGlnGlyValaGlyLeuAlaAspGluTrpAsnAlaAspIle 251
Db 541 TTTCAATTTAATTTATTAACAACAAGTCTGATTTGGCTGATGAAATGCAATATCA 600
QY 252 HisProSerGlnIleGluProAsnAlaGlyThrSerAspAspTyTyIlyLeuLeuIlyS 271
Db 601 CATCCTTCACAAAATTTGAACCTTAATGCTGGAACATCAGATGACTATTATAAATTTTAAAA 660
QY 272 GluAsnIleProIlyTyThrSerAsnTyrcyAlaIleAsnThrTyraGThGlyLeuIlyAsn 291
Db 661 GAAATATATACCTTAATATATAGTACTATTTGCAATATCTTAATAGAGAGACTAAATATA 720
QY 292 LeuArgAspGluProAsnMetIlyTrpSerIlePheAsnAspTyArgArgTyrcMetThr 311
Db 721 CTTGGAACGAACTTAATATATAGATGAGTATTTATTAATGATTTATCGAAGTATATGACT 780
QY 312 ILeuThrValLeuAspThrIleSerGlnPheSerleuTyraAspIleIlyArgTyArgAsp 331
Db 781 ATTAAGTATTAAGTACTATCGCTCAATTTCTTTTATGATATTAAGATATACAAAGAT 840
QY 332 SerIleGlyGlyIleGlyValaIlyGlyIlyAsnGlnleuThrArgGluIleTyThr 351
Db 841 TCAATATAGA-----AGATATGTTGGCATTAACCTGAACCTTAACAGAAATTTATACA 894

```

Oy	352	ThrGluIleAsnPhleAspArgLeuProGlnLeuArgValGlnProAsnLeuAlaThrMet	372
Db	895	ACTGAAATAAATTTTGGACCGCTTACTTACCTTGAAATTCACCACTTCGCTATATATG	954
Oy	372	GluTyArgLeuLeuThrArgAlaSerPheLysLeuPheSerPheLeuGluGlnPheIlePhe	391
Db	955	GAAATATATTTTAAACGCTTCAGGGCTTAGATTATTTTCATTTTTCATTTTTCATTTTTC	1014
Oy	392	TyrThrGluAsnThrAsnPheGlyAsnArgLeuValGlyIleSerAsnArgAspAlaPro	411
Db	1015	TATACAAAATAATGAAACGTACGGGAATCGTTTAGTTGGTATTCGGAAATCGTAAATGAT	1074
Oy	412	ThrTySerAsnThrIleThrGluThrLeuTyArgIleGluAlaGlnThrGlySerProThrThr	431
Db	1075	ACTTATGCTTACGACAGCAAGAACGTAAATTATATATGAGAAAGAACAGGCCACCAACA	1134
Oy	432	LysThrIleAsnGProPheGluSerTyArgValSerIleValThrAspArgGlnSerPro	451
Db	1135	AAACCTTAAATACCATTTTGAATTCCTATTAAGCTTCAATGTACGTATGACCACTACT	1194
Oy	452	ProValSerProIle---GlnProIlePheIleIleAsnGlnIleGluLeuTyArgAsn	470
Db	1195	CCATCTTCCCTTTTCCATACATATACTTTCACATTAATCAATTAATGAACCTTATTTAAAT	1254
Oy	471	GlySerSerAsnAsnThrLeuLysTySerAlaGlyLysLeuSerLeuSerAsnTyArgAsn	490
Db	1255	AATTCACCTAGTAAATTAATTAACATATATCACTGGGGGGAATTTATCTTAATGATTA	1314
Oy	491	ThrThrPhePheGlnPheProArgLysLysArgCysAsnLeuValIleAspProGlyCys	510
Db	1315	ACAACTGATTTTCAATTTCCGTAAAAAACAACGTAAACCAATTAATTAATCCAAATGT	1374
Oy	511	SerProAsnPheAsnAsnTySerHisIleLeuSerHisPheSerLeuPheThrTySer	530
Db	1375	TTACCAAGCTATTAATGATTAATGATCAATTTTATCCAGTTTCTTATTTAATTAATTC	1434
Oy	531	TyrValIleGlyLeuGlnLeuGlnIleLeuAspThrGlyValLeuGlyTyrThrHisSer	550
Db	1435	TATATAATTTGATTAAGCGCTAAATATATATATATACAGGTGATTAAGATGACACACGT	1494
Oy	551	SerValAspArgTyArgAsnAlaIleSerAspLysIleIleThrMetIleProAlaIleLys	570
Db	1495	AGGTTTATATACAAATAATGCAATTCAGATTAATAATTAACAAATGATCCCGACATCMA	1554
Oy	571	GlyAsnAsnLeuAspThrAsnSerLysValIleGluGlyProGlyHisThrGlyGlyAsn	590
Db	1555	GGTAAACGCTTGTGANTACCAACTCTTAAGGTAATTGAAGACCTGGTCATACAGAGAAAC	1614
Oy	591	LeuValTyArgLeuGlnSerGlnTyArgLeuGlnIleThrCysGluThrProAsnSerThr	610
Db	1615	TTGGTTATTTACAAAGTCAAGGGCGTTTAAGATTACATGTAGAACCTCCATATTCTACA	1674
Oy	611	GlnSerTyPheIleArgLeuArgTyValaThrArgIleValaGlyAsnThrLeuProAsn	630
Db	1675	CAATCTTATTCATTAGACTTCGATTCGATACGCTCAAAATGGTGTGGAATAACTCTTCAT	1734
Oy	631	IleSerLeuThrIleProGlyValIleGlyIleProProGlnArgLeuAsnAsnThrPhe	650
Db	1735	ATATCTCTTACAAATACGAGGATATAGAAATACCACTTCACCACTCAACAAACATTTT	1794
Oy	651	SerGlyThrAsnTyArgAsnLeuGlnTyArgLysPheGlyTyArgPheGlnPheProSer	670
Db	1795	TCTGGTACAAATTAATTAATTTTACAAATCGAGATTTTGGGATTTTCCAAATTTCCAGT	1854
Oy	671	ThrValThrLeuProLeuAsnArgAsnIleProPheIlePheAsnArgAlaAspValSer	690
Db	1855	ACAGTAACATTAACCTTTAAATCGAAACATCACTATTATTTAATCGTACAGATGATACA	1914
Oy	691	AsnSerIleLeuIleIleAspLysIleGluThrIleProIleThrSerSerMetHisGln	710
Db	1915	AATTCAAATTTAATCATGATTAATTAATTAATTTAATCAAAATTAATCTCTCTGACGCCA	1974

```

QY      711  AanaagGluLysGlnLysLeuGlnThrIleGlnThrLysIleAsnTrpPheThrsn  730
          |||||
Db       1975  AATAGAGAAAAACAAATAATTAGAACTATCCAAACAAAATAAATACATTTTTCACAAAT  2034
          |||||

QY      731  HistHrLys 733
          |||||
Db       2035  CATCAAAA 2043

RESULT 4
US-09-756-526A-3
; Sequence 3, Application US/09756526A
; Patent No. US20020038005A1
GENERAL INFORMATION:
APPLICANT: Jana, Wojciechowska
APPLICANT: Evgeny, Lewitin
APPLICANT: Ludmila, Revina
APPLICANT: Igor, Zalunin
APPLICANT: Galina, Chestukhina
TITLE OF INVENTION: NOVEL DELTA-ENDOTOXINS AND NUCLEIC ACID SEQUENCES CODING THEREFOR
FILE REFERENCE: S-30913A
CURRENT FILING DATE: 2001-01-08
PRIOR APPLICATION NUMBER: US 60/175,158
NUMBER OF SEQ ID NOS: 5
SOFTWARE: PatentIn version 3.0
SEQ ID NO 3
LENGTH: 4896
TYPE: DNA
ORGANISM: Bacillus thuringiensis
FEATURE:
NAME/KEY: source
LOCATION: (1)..(4896)
OTHER INFORMATION: Bacillus thuringiensis supsp. finitimus strain VKPM B-1161 (taxon
OTHER INFORMATION: :29337
NAME/KEY: CDS
LOCATION: (1129)..(4458)
OTHER INFORMATION: product: Cry28Aa1 delta-endotoxin
PUBLICATION INFORMATION:
AUTHORS: Wojciechowska, et al.
TITLE: Two novel delta-endotoxin gene families cry26 and cry28 from Bacillus
JOURNAL: FEBS Lett.
VOLUME: 453
ISSUE: 12
PAGES: 46-48
DATE: 1999-06-18
DATABASE ENTRY DATE: ____-__-__
US-09-756-526A-3

Alignment Scores:
Pred. No.: 1,54e-79 Length: 4896
Score: 849.00 Matches: 229
Percent Similarity: 46.8% Conservative: 106
Best Local Similarity: 33.3% Mismatches: 248
Query Match: 21.9% Indels: 104
DB: 3 Gaps: 25

US-10-783-417-2 (1-735) x US-09-756-526A-3 (1-4896)
QY      68  SeralaasPrHlleAlaValSerAlaGlyThrIleValSerGlyThrLeuLeuAla  87
          ||::|||:|||||::|||::|||
Db       1210  AGCAGTAGATACAGTGCGCTGTAGTAAGCGGAGGATGTAGTGGGGACTATACTGACA  1269
          |||||

QY      88  GlyIleGlyLeuThrSerIleSerGlyProIleGlyIleIleGlyAlaIleIle  107
          |||||
Db       1270  -----GCCTTGATCATTTGTGAATCCA-----GGGTGGTACTTATTA  1308
          |||||

QY      108  SerPheGlyThrLeuIleThrValPheTrpProAlaGlyGluGlnAsp---LysThrVal  126
          |||||
Db       1309  TCAITTTGAACTTGGCTGCCGTTCTTTGGCTTGATCCAGAAGAACATCCAAAAAATT  1368
          |||||

QY      127  TrpHrgInPheIleLysMetGlyGlnIlePheValAspThrProLeuThrGluSerIle  146

```

Db	1369	TGTCACAACTTATATGAAACCGGACAAACCTTTAAATCAAAACAATTTCTACAGCTGTAA	1422
Oy	147	LygclnleuLysleuGlnThrleuGlnGlyPheArgGlnlleuGlnSerTyraThr	166
Db	1429	AAAGAAATGCAATTGACTCATCTTAATGCTTTTAAAGATGTATTAACGTACTATCAAGA	1488
Oy	167	AlaIeuAAspTrpArgIleuLysleuLysArgleuGlnAlaProGlyLeuProSerSer	186
Db	1489	GCATTATATGATTTGAAGAGA-----AATCAAGATGCA	1522
Oy	187	AlaIeuGlnGlnAlaIleuThrleuLyslleArgPheGluAsnValHisAsnAspPhe	206
Db	1522	-----AATACGCCAGATGGTATACACAGACATTTGAAAAGCTCATTTCAATTTT	1572
Oy	207	lleArgGlnlleProGlyPheGlnleuGlnThrTyLysThrLeuLeuLeuProIleTy	226
Db	1573	GTAAGCAATATGCCCAACCACTCCCACTCCACGATGACACACTTTATTATTAAGTGTAT	1632
Oy	227	AlaGlnAlaAlaAsnPheHisleuAsnleuGlnGlnGlnAlaGlnLeuAlaAspGlu	246
Db	1633	ACAGAAAGCTGCMAATTACATTGTAATTATTAACATCAAGGGTACAAATTCGGGATCAA	1692
Oy	247	TrpAsnAlaAspIleHisProSerGlnlleGluProAsnAlaGlyThrSerAspArgTy	266
Db	1633	TGGATGCGAGATCAACCAACATTCACCAATGTGAAGTACATCAGACT-----TAT	1742
Oy	267	TyrIlySleuLeuLysGlnAsnIleProLysTySerAsnTyTyAlaAsnThrTyArg	286
Db	1744	TATGACGAGCTATGGTATATGTATTAAGTAAAGTATTAATTAATGACCAACAGACATACAT	1802
Oy	287	ThrGlyLeuLysAsnLeuArgAspGluProAsnMetLysTrpSerIlePheAsnAspTy	306
Db	1804	AAAGGATTTGAATCACCTTAACAAACACAGAAAAATCACATGGGATGCTTAATACACATAT	1862
Oy	307	ArgArgTyMetThrIleThrValLeuAspThrIleSerGlnPheSerIleuTyraPile	326
Db	1864	CGTCAGAAATACACTTATATTTGATTTGATCTTTCGCAACTTTTCCTTTTATATATA	1922
Oy	327	LysArgTyArgAspSerIleGlyGlylleGluValIleGlyAsnGlnLeuThr	346
Db	1924	CGTCGTTT-----CCAGAGGAGTGAACCTACATTAATTAACA	1952
Oy	347	ArgGlnIleTyThrThrGlnIleAsnPheAspArgLeuProGlnleuArgValGlnPro	366
Db	1960	AGAGGGGTTTATACAGATTAGATCATTTTAAACACACACCA-----	2002
Oy	367	AsnIeuAlaThrMetGluTyraSnuLeuThrArgAlaSerPheLysLeuPheSerPheLeu	386
Db	2002	-----GGGCTATTATTCCTGGCTG	2012
Oy	387	GluGlnPheIlePheTyThrGlnAsnThrAsnPheGlyAsnArgLeuValGlylleSer	406
Db	2020	TCAGATATTAGGTATATACACGAGAGGTGGCAGAAAGGAGATTAATTATACAGTATTT--	2076
Oy	407	AsnArgAspAlaProThrTySerAsnThr-----lleThrGlnThrleuTyrgly	423
Db	2077	---CGAGAGCTCAATATATATCTGTATCAATTTTTCAGATCAAAAATATTTATAGGT	2132
Oy	424	GluArgThrGly---SerProThrThrIysThrIleArgProPheGlnSerTyrylVal	442
Db	2134	AATACAAATAGATTAAAGTAAAGCAGCTCATTTACATTAATTACACAGGCGAA--TTATAGACT	2196
Oy	443	SerIleValThrAspArgGlnSerProProValSerProIleGlnProHisIlellele	462
Db	2191	CACCTTAAGATTAACCGTCTCTTTCAAAACAATAGCTCGATATAAATAGATTAAACGTTTA	2252
Oy	463	AsnGlnIleGlnLeuTyrylLeuAsnGlySerSerAsnAsnThrLeuIysTySerAlaGly	482
Db	2251	ATTCAAAAATATCGTATTCACAACTTTTAAAAACGATTAAT-----GAATATCAAAAAAAT	2304
Oy	483	GlySerSerSerAsnTyTy-----GlnAsnThrThrPhePheGlnPheProArgLys	499

```

Db      2305 TTTAATGTGAATTAATCAAAATGAACCTCAAGAAACTACA----- 2343
Qy      500 IySaSpCyAsnLeuValIleAspProGlyCySerProAsnPhaAsnTyrSerHis 519
      |||
Db      2344 -----AACTATCTTAATGATTAATGAGTGTCA--AACACCCAAAAATTCAAACT 2391
      |||
Qy      520 IleuSerHisPheSerLeuPheThrTyrSerTyrValIleGlyLeuGlnLeuGlnIle 539
      |||||
Db      2392 AATTATCTCATTTTCATTATCATCCAC-----AAGTTAGAGTTT 2433
      |||
Qy      540 LeuAsp-----ThrGlyValLeuGlyTyrThrHisSerSerValAspArgTyr 555
      |||
Db      2434 GCGAGATATTTTCATCTATATTTGATAGTTAGGTGACACACAAATAGTAACTCCCA 2493
      |||
Qy      556 AenAlaIleSerAspIleIleThrMetCileProAlaIleIleGlyAsnAsnLeuAsp 575
      |||
Db      2494 AATTTAATTCAGAAAAGTGTAGTACCAAAATCCCATTTGTAAGCTTACCAAGTT--- 2550
      |||
Qy      576 ThrAsnSerIleValIleGlyGlyProGlyHisThrGlyGlyAsnLeuValTyrLeuGln 595
      |||||
Db      2551 ACTAACAAATTCAGTTTAAAGAGACACAGGTTTACAGGTGAGAGATTTAATGAACCTTGT 2610
      |||||
Qy      596 SerGlnGlyArgLeuGluIleThrCysGluThrProAsnSerThrGlnSerTyrPheIle 615
      |||
Db      2611 GAT-----AAATGTTCTATTAATGTAAGCT--AGTTCTTAAAAAATAGCTATA 2661
      |||
Qy      616 ArgLeuArgTyrAlaThrAsnGlyAlaGlyAsnThrLeuProAsnIleSerLeuThrIle 635
      |||
Db      2662 AGCTCAATTTATGCTCCAAATTAACGAAATAGCTGATCAATTAAGACGTGATTCGGA 2721
      |||||
Qy      636 ProGlyValIleGlyIleProGlnArgLeuAsnAsnThrPheSerGlyTyrAsnTyr 655
      |||||
Db      2722 GCAGGAGTTCTA-----TTGCACCTTACCTTTTCTAGAAAAGGAGAAC 2763
      |||||
Qy      656 AsnAsn-----LeuGlnTyrGlyAspPheGlyTyrPheGlnPheProSer 670
      |||||
Db      2764 AATTAATTTACATTCAGACGCTTAATAGAGATTTTCATATCATACACTTTAGTT 2823
      |||||
Qy      671 ThrValThrLeuProLeuAsnArgAsnIleProPheIlePheAsnArgAlaAsp----- 688
      |||
Db      2824 GATATTGAATTTACCCCAAGGTGAAGAAATTCATATTCATTGAAAGCGAGAGATCATTTAT 2883
      |||
Qy      689 ValSerAsnSerIleLeuIleIleAspIleIleGluPheIleProIleThrSerSerMet 708
      |||||
Db      2884 GAGGAGGAGGTGATCTTTTAATTTGATTAATTAAGAGTCAAACTATAGATGAATAATAT 2943
      |||||
Qy      709 HisGlnAsnArgGluIleGlnIleGlnIleGlnIleGlnIleGlnIleGlnIleGlnIle 728
      |||
Db      2944 -----ACTATATGAATGAATTTAGAGAGCAAGCAAGAAAGCAAGTATATATTT 2994
      |||
Qy      729 ThrAsnHisThrIleThrLeu 735
      |||
Db      2995 ATTAACGCAACAACGCTTGG 3015
      |||

RESULT 5
US-10-345-020-3
/ Sequence 3, Application US/10345020
/ Publication No. US20030150018A1
/ GENERAL INFORMATION:
/ APPLICANT: Jana, Wojciechowska
/ APPLICANT: Evgeny, Lewitin
/ APPLICANT: Ludmila, Revina
/ APPLICANT: Igor, Zalumin
/ APPLICANT: Galina, Chestubhina
/ TITLE OF INVENTION: NOVEL DELTA-ENDOTOXINS AND NUCLEIC ACID SEQUENCES CODING THEREFOR
/ FILE REFERENCE: S-309113
/ CURRENT APPLICATION NUMBER: US/10/345,020
/ PRIOR FILING DATE: 2003-01-15
/ PRIOR APPLICATION NUMBER: US 60/175,158
/ PRIORITY FILING DATE: 2000-01-07
/ NUMBER OF SEQ ID NOS: 5
/ SOFTWARE: PatentIn version 3.0
/ SEQ ID NO 3

```

```

/ LENGTH: 4896
/ TYPE: DNA
/ ORGANISM: Bacillus thuringiensis
/ FEATURE:
/ NAME/KEY: source
/ LOCATION: (1)..(4896)
/ OTHER INFORMATION: Bacillus thuringiensis subsp. finitimus strain VKPM B-1161 (taxon
/ OTHER INFORMATION: 29337
/ FEATURE:
/ NAME/KEY: CDS
/ LOCATION: (1129)..(4458)
/ OTHER INFORMATION: product: Cry28Aa1 delta-endotoxin
/ PUBLICATION INFORMATION:
/ AUTHORS: Wojciechowska, et al.
/ TITLE: Two novel delta-endotoxin gene families cry26 and cry28 from Bacillus
/ TITLE: thuringiensis ssp. finitimus
/ JOURNAL: FEBS Lett.
/ VOLUME: 453
/ ISSUE: 12
/ PAGES: 46-48
/ DATE: 1999-06-18
/ US-10-345-020-3
/ US-10-783-417-2 (1-735) x US-10-345-020-3 (1-4896)

```

```

Alignment Scores:
Pred. No.: 1,54e-79 Length: 4896
Score: 849.00 Matches: 229
Percent Similarity: 48.8% Conservative: 106
Best Local Similarity: 33.3% Mismatches: 248
Query Match: 21.9% Indels: 104
Gaps: 25

```

```

US-10-783-417-2 (1-735) x US-10-345-020-3 (1-4896)
QY 68 SerAlaAPRThrIleAlaAlaValSerAlaGlyThrIleValSerGlyThrLeuAla 87
Db 1210 AGCAGTATACAGTGTGCTGTAGTAAGCGAGGATGTTGCTGTACTATACAGCA 1269
QY 88 GlyIleGlyLeuThrSerIleSerGlyProIleGlyIleIleGlyAlaIleIle 107
Db 1270 -----GCCTTGACATCATTTGTAATCCA-----GGTGTGACTATA 1308
QY 108 SerPheGlyThrLeuIleThrValPheTrpProAlaGlyGluGlnAsp---LysThrVal 126
Db 1309 TCATTGGAACTTGCTGCTCCGCTTCTTGCTGATCCAGAGAAATCCAAAAAATT 1368
QY 127 TrpThrGlnPheIleLysMetGlyGluIlePheValAspThrProLeuThrGluSerIle 146
Db 1369 TGGTCACATTTATGAAACACGAGAGAACCTTTTAATCAACAATTTCTACAGCTGA 1428
QY 147 LysGlnLeuLysLeuGlnThrLeuGluGlyPheAspGlnIleLeuGlnSerTyrrhenth 166
Db 1429 AAGAAATGATGATTCATCATCTTAATGTTTAAAGATGATTAACGTACTATGAAAGA 1488
QY 167 AlaLeuAspAspTrpArgLysLeuLysArgLeuGlnAlaProGlyLeuProProSerSer 186
Db 1489 GCATTTAATGATTTGAAAGAA-----AATCCAGTGA 1521
QY 187 AlaLeuGlnAlaAlaLeuThrLeuLysValIleArgPheGluAsnValHisAsnAspPhe 206
Db 1522 -----AATACGCGCAGATTTGGTATCACAAGATTTGAAAGCCTCATTTCAATTTT 1572
QY 207 IleArgGluIleProGlyPheGlnLeuGluThrTyrrLysThrLeuLeuLeuProIleTy 226
Db 1573 GAAAGCAATATGCAAACTCCAACTCCAACTCCAGCTATGACACATTTATTAAGTTGCTAT 1632
QY 227 AlaGlnAlaAlaAsnPheHisLeuAsnLeuLeuGlnGlnGlyValaGluLeuAlaAspGlu 246
Db 1633 AAGAAAGCTGCAAATTTACATTTGATTTATTTACATCAAGGTGTACAAATCCGAGATCA 1692
QY 247 TrpAsnAlaAspIleHisProSerGlnIleGluProAsnAlaGlyThrSerAspAspTy 266
Db 1693 TGGAAATGACATCAACCAATTCACCAATGTTGAAGTCAATGCTACT-----TAT 1743

```

```

QY 267 TyrIleLeuLeuLysGluAsnIleProLysTyrrSerAsnTyrrCysAlaAsnThrTyrrArg 286
Db 1744 TATGACGAGCTATGCTATATATTTGAAAAGTATATTAATTTATTCACCAAGCATCATCAT 1803
QY 287 ThrGlyLeuLysAsnLeuArgAspGluProAsnMetLysTrpSerIlePheAsnAspTy 306
Db 1804 AAGGATGTAATTCACCTTAAGAAATCAGAAAAAATCACAATGCAATGCTATATACATAT 1863
QY 307 ArgArgTyrrMetThrIleThrValLeuAspThrIleSerGlnPheSerLeuTyrrAspIle 326
Db 1864 GCTCGAAGAAATGACCTTAATGTATGTATGATCTTGCGCACTTTCCTTTATATATATA 1923
QY 327 LysArgTyrrArgAspSerIleGlyIleGluValLysGlyIleLysAsnGluLeuThr 346
Db 1924 CCGCTTTT-----CCAAAGAGATGTAAGTAAATTAACA 1959
QY 347 ArgGluIleTyrrThrGluIleAsnPheAspArgLeuProGlnLeuArgValGlnPro 366
Db 1960 AGAGAGATTATACAAATTAGATCATTTAACACGACCA----- 2001
QY 367 AsnLeuAlaThrMetGluTyrrAsnLeuThrArgAlaSerPheLysLeuPheSerPheLeu 386
Db 2002 -----GGCGCTATTTACTTGCTG 2019
QY 387 GluGlnPheIlePheTyrrThrGluAsnThrAsnPheGlyAsnArgLeuValIleSer 406
Db 2020 TCAAGATTTAGATTTATACAGAGAGTGTGCGAAGCGATTTATTTACAGTATTT 2076
QY 407 AsnArgAspAlaProThrTyrrSerAsnThr-----IleThrGluThrLeuTyrrGly 423
Db 2077 ---CGAGAGTCAATTAATTAATTAATCTGTAATCAATTTTTCAGATGAAAAAATTTATATGCT 2133
QY 424 GluArgThrGly---SerProThrTyrrSerThrIleAspProPheGluSerTyrrLysVal 442
Db 2134 AATACAAATAGATTAAAGTAAGCAGCTCATTTATTAATTAACAGGCGAA---TTATAGCT 2190
QY 443 SerIleValThrAspArgGlnSerProProValSerProIleGlnProHisPheIleIle 462
Db 2191 CACTTAAGCATTAACCGTCTTTTCAACAACATGCTGGATTAATAAGTTATACACTTTA 2250
QY 463 AsnGlnIleGluLeuLysAsnGlySerSerAsnAsnThrLeuLysTyrrSerAlaGly 482
Db 2251 ATTCAAAAATGTAATTCACAACCTTTAAACGATAT-----GAATATCAAAAAAAT 2304
QY 483 GlySerLeuSerAsnTyrr-----GlnAsnThrThrPhePheGlnPheProArgLys 499
Db 2305 TTTAATGTAATTAATCAAAAATGAACTCAAGAAACTTACA----- 2343
QY 500 LysAspCysAsnLeuValIleAspProGlyCysSerProAsnPheAsnAsnTyrrSerHis 519
Db 2344 -----AATATCTTAATATTAATGATGCTCA---AACAGCCAAAATTCAAAACAT 2391
QY 520 IleLeuSerHisPheSerLeuPheThrTyrrSerTyrrValIleGlyLeuGlnGlnIle 539
Db 2392 AATTTATCTCATTTTCCATTAATCATCAAC-----AAGTTAAGGTTT 2433
QY 540 LeuAsp-----ThrGlyValLeuGlyTyrrThrHisSerSerValAspArgTy 555
Db 2434 GCTGATATTTTCACTATATTTTCACTATGAGTGTGACACACAATATGTTAACTCCCA 2493
QY 556 AsnAlaIleSerAspLysIleIleThrMetIleProAlaIleLysGlyAsnAsnLeuAsp 575
Db 2494 AATTTAATATCAAGAAAGTGTGATCAAAATCCCTTGCTAAAGCTTACAGAAATT--- 2550
QY 576 ThrAsnSerLysValIleGluGlyProGlyHisThrGlyGlyAsnLeuValTyrrLeuGln 595
Db 2551 ACTTAACAATTCAGTTATTAAGAGACAGAGTTTACAGGAGGATTTAATACAACTTCCT 2610
QY 596 SerGlnGlyArgLeuGluIleThrCysGluThrProAsnSerThrGlnSerTyrrPheIle 615
Db 2611 GAT-----AATGTTCCTATTAATGTAAGCT---AGTTCTTTAAAAAATACCTATA 2661
QY 616 ArgLeuArgTyrrAlaThrAsnGlyValaGlyAsnThrLeuProAsnIleSerLeuThrIle 635

```

```

Db      2662 AGCTATTATTGCTGCAAAATACGCAATAGCTGATCATATAGACGTAGTGGATCCGGA 2721
Qy      636 ProGlyValIleGlyIleProProGlnArgLeuAenAenThrPheSerGlyThrAsnThr 655
Db      2722 GCAGAGAGTTCTA-----TTGCAACTACTCTTTTCTTGAAAGAGGAAAC 2763
Qy      656 AenAen-----LeuGlnIyrGlyAepPheGlyTyrPheGlnPheProSer 670
Db      2764 AATAATTTTACAAATCAAGACCTTAATAGATTTCATATTCATATCATACCTTTAGTT 2823
Qy      671 ThrValThrLeuProLeuAenArgAsnIleProPheIlePheAenArgAlaAsp----- 688
Db      2824 GATATGTAATTCACGAAATGAGAAATTCATATCCATTTGAAACGAGAGATGATTAT 2883
Qy      689 ValSerAenSerIleLeuIleIleAspIyrIleGluPheIleProIleThrSerSerMet 708
Db      2884 GAGGAGGAGATGATCTTTTATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2943
Qy      709 HisGlnAenArgGluIleGlnIleGluIleGluIleGlnIleGlnIleGlnIleGlnIle 728
Db      2944 -----ACTAATGAATGAATTTAGAGAGGAAAGAAAGAGTGAATGATATT 2994
Qy      729 ThrAenHisThrIleThrLeu 735
Db      2995 ATAAACGACAAACGCTTTG 3015

RESULT 6
US-10-342-821-3
; Sequence 3, Application US/10342821
; Publication No. US20030154510A1
; GENERAL INFORMATION:
; APPLICANT: Jana, Wojciechowska
; APPLICANT: Evgeny, Lewitin
; APPLICANT: Ludmila, Revina
; APPLICANT: Igor, Zalunin
; APPLICANT: Galina, Chesukhina
; TITLE OF INVENTION: NOVEL DELTA-ENDOTOXINS AND NUCLEIC ACID SEQUENCES CODING THEREFOR
; FILE REFERENCE: S-30913C
; CURRENT APPLICATION NUMBER: US/10/342, 821
; PRIOR FILING DATE: 2003-01-15
; PRIOR APPLICATION NUMBER: US 60/175,158
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: Patent in version 3.0
; SEQ ID NO 3
; LENGTH: 4896
; TYPE: DNA
; ORGANISM: Bacillus thuringiensis
; FEATURE:
; NAME/KEY: source
; LOCATION: (1)..(4896)
; OTHER INFORMATION: Bacillus thuringiensis supsp. finitimus strain VKPM B-1161 (taxon
; OTHER INFORMATION: :29337
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1129)..(4458)
; OTHER INFORMATION: product: Cry28a1 delta-endotoxin
; PUBLICATION INFORMATION:
; AUTHORS: Wojciechowska, et al.
; TITLE: Two novel delta-endotoxin gene families cry26 and cry28 from Bacillus
; JOURNAL: FEBS Lett.
; VOLUME: 453
; ISSUE: 12
; PAGES: 46-48
; DATE: 1999-06-18
US-10-342-821-3

Alignment Scores:
Pred. No.: 1,54e-79 Length: 4896
Score: 849.00 Matches: 229
Percent Similarity: 46.8% Conservative: 106

```

```

Best Local Similarity: 33.3% Mismatches: 248
Query Match: 21.9% Indels: 104
DB: 6 Gaps: 25

US-10-783-417-2 (1-735) x US-10-342-821-3 (1-4896)
Qy      68 SerAlaAspThrIleAlaIleValSerAlaGlyThrIleValSerGlyThrLeuLeuAla 87
Db      1210 AGCAGTGTATACGTCGCTGTAGTAAGCCAGAGGATGTAGTGTGGGTACTATATCTGCA 1269
Qy      88 GlyIleGlyLeuThrSerIleSerGlyProIleGlyIleIleGlyAlaIleIleIle 107
Db      1270 -----GCCTTGACATCTTTGTTATCA-----GGTGTGACTTATTA 1308
Qy      108 SerPheGlyThrLeuIleThrValPheTyrProAlaGlyGlnGlnAsp-----LysThrVal 126
Db      1309 TCATTGGAACTTGCGCTCGGCTCTTGGCTGATCCAGAGAAATCCAAAATAATT 1368
Qy      127 TrpThrGlnPheIleLysMetGlyGluIlePheValAspThrProLeuThrGluSerIle 146
Db      1369 TGGTCACATTTATGAAACACGAGAAAGACCTTTAATCAAAATTTCTACAGCTGTA 1428
Qy      147 LysGlnLeuIleGlnIleGlnIleGlnIleGlnIleGlnIleGlnIleGlnIleGlnIle 166
Db      1429 AAAGAAATAGCAATTCATCTCAATTAATGTTTAAAGTGTATTAACCTTACTATGAAGA 1488
Qy      167 AlaLeuAspAspTyrPargLysLeuIleValArgLeuGlnAlaProGlyLeuProProSerSer 186
Db      1489 GCATTATTAATGTTGGAAGA-----AATCCAAGTCA 1521
Qy      187 AlaLeuGlnGlnAlaIleLeuThrLeuIleLysIleArgPheGluAenValHisAenAspPhe 206
Db      1522 -----AATACTGCGAGATGGTATCACAGAGATTGAAAACGCTCATTTCAATTT 1572
Qy      207 IleArgGluIleProGlyPheGlnLeuGlnIleThrIleIleIleIleIleIleIleIleIle 226
Db      1573 GTAAGCAATATGCCCAACTCCACTCCAGTGTACACATTAATTAATTAATGTTGCTAT 1632
Qy      227 AlaGlnAlaIleAenPheHisLeuAenLeuLeuGlnGlnIleGlyAlaGluLeuAlaAspGlu 246
Db      1633 ACAGAACTGCCAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1692
Qy      247 TrpAenAlaAspIleHisProSerGlnIleGluProAenAlaGlyThrSerAspAspTyr 266
Db      1693 TGGAAATGACATCAACACATTCACCAATGTTGAAGTCATCAGGTAATCT-----TAT 1743
Qy      267 TyrIleLeuLeuIleGluAenIleProLysTyrSerAsnTyrCysAlaAsnThrTyrArg 286
Db      1744 TATGACGAGCTATTGTTATATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1803
Qy      287 ThrGlyLeuLysAenLeuArgAspGluProAenMetLysTyrSerIlePheAenAspTyr 306
Db      1804 AAAGATTAATACACTTAAAGAAATCAGAAAATACATGGAGTCTTATTAACCATAT 1863
Qy      307 ArgArgTyrMetThrIleThrValIleLeuAspThrIleSerGlnPheSerLeuTyrAspIle 326
Db      1864 COTGAGAAATGACCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1923
Qy      327 LysArgTyrArgAspSerIleGlyIleGluValIleGlyIleLysAsnGluLeuThr 346
Db      1924 CGTCGTTT-----CCAAAGAGAGTGAAGCTAGAAATTAACA 1959
Qy      347 ArgGluIleTyrThrThrGluIleAenPheAspArgLeuProGlnLeuArgValGlnPro 366
Db      1960 AGAGAGTTTATACAGTTTATGATTAATTAACACACACCA----- 2001
Qy      367 AsnLeuAlaThrMetGluTyrAsnLeuThrArgAlaSerPheLysLeuPheSerPheLeu 386
Db      2002 -----GGGCTATTACTTGCTG 2019
Qy      387 GluIlePheIlePheTyrThrGluAenThrAenPheGlyAenArgLeuValGlyIleSer 406
Db      2020 TCAGATATTGATTATACACGAGAGTGTGGCAGAAAGCATTAATTAATTAATTAATTAATTA 2076

```

407 AsnArgAspAlaProThrTyrSerAsnThr-----IleThrGluThrLeuTyrGly 423
|||:|||||
2077 ---CGAGAGCTAAATATATATACGTGATATCAATTTTTCACATGAAAAATTTATAGCT 2133
424 GluArgThrGly---SerProThrThrIleArgProPheGluSerTyrIleVal 442
|||:|||||
2134 AATACAAATAGATTAAGTAAAGCAGCTCATCATTAATTCACGGGGA---TTTATGACT 2190
443 SerIleValThrAspArgGlnSerProProValSerProIleGlnProHisPheIleIle 462
|||:|||||
2191 CACTTAAGCATTAACCGCTCTTTTCACAAATAGCGGTATTAATTAAGTTAATACAGTTTA 2250
463 AsnGlnIleGluLeuThrLeuAsnGlySerSerAsnAsnThrLeuIleTyrSerIleGly 482
|||:|||||
2251 ATTCAAAAAATGATATTCACAACTTTTAAAAACGATAT-----GAATATCAAAAAAAT 2304
483 GlySerLeuSerAsnTyr-----GlnAsnThrThrPhePheGlnPheProArgIle 499
|||:|||||
2305 TTTAATGTGAATATATCAAAATGACCTCAAGAACTACG----- 2343
500 LysAspCysAsnLeuValIleAspProGlyCysSerProAsnPheAsnAsnTyrSerHis 519
|||:|||||
2344 -----AACTATCCTAATGATTAAGTGGTTCG---AACAGCCAAAAATTCAAACAT 2391
520 IleLeuSerHisPheSerLeuPheThrTyrSerTyrValIleGlyLeuGlnLeuGlnIle 539
|||:|||||
2392 AATTAATCATATTTCCATTAATCATCCG-----AAGTTAGAGTTT 2433
540 LeuAsp-----ThrGlyValLeuGlyTyrThrHisSerSerValAspArgTyr 555
|||:|||||
2434 GGTGAGATTTTTCACCTCATATTTGATAGTGGACACACATAGTATAAGTCCCA 2493
556 AsnAlaIleSerAspIleIleIleThrMetIleProAlaIleIleGlyAsnAsnLeuAsp 575
|||:|||||
2494 AATTAATATTCAGAAAGTGTGAGTACACAAATCCCATGTGTAAGCTTACAGAGTT--- 2550
576 ThrAsnSerLeuValIleGluGlyProGlyHisThrIleGlyAsnLeuValTyrLeuGln 595
|||:|||||
2551 ACTTAACATTCAGTTATTAAGAGACCAAGGTTTTCAGGAGATTTAATTAACCTTCCT 2610
596 SerGlnGlyArgLeuGluIleThrCysGluThrProAsnSerThrGlnSerTyrPheIle 615
|||:|||||
2611 GAT-----AAATGTTCTTATTAATGTAAGCT---AGTTCTTTAAAAAAATACCTATA 2661
616 ArgLeuArgTyrAlaThrAsnGlyIleGlyAsnThrLeuProAsnIleSerLeuThrIle 635
|||:|||||
2662 AGTCAATTTTATGCTGCAAAATAGCAATAGCTGATCAATAGACGTAGTATCCGGA 2721
636 ProGlyValIleGlyIleProProGlnArgLeuAsnAsnThrPheSerGlyThrAsnTyr 655
|||:|||||
2722 GCAGGAGTCTA-----TTGCACCTTACTTTTTCAGAAAGGGAAC 2763
656 AsnAsn-----LeuGlnTyrGlyAspPheGlyTyrPheGlnPheProSer 670
|||:|||||
2764 AATATTTTTCACATTCAGACCTTAACATAAGATTTTTCATATCATACCTTTAGTT 2823
671 ThrValThrLeuProLeuAsnArgAsnIleProPheIlePheAsnArgAlaAsp----- 688
|||:|||||
2824 GATATTTGATTAACCGAAGTGAAGAAATTCATATTCATTTGACGAGAGATGATAT 2883
689 ValSerAsnSerIleLeuIleIleAspIleIleGluPheIleProIleThrSerMet 708
|||:|||||
2884 GAGGAGGAGTATCTTTTAATGATTAATTAAGGTTCAACCTATTAATGAAAAATTA 2943
709 HisGlnAsnArgGluGlyGlnIleGluGluThrIleGlnThrIleAsnThrPhePhe 728
|||:|||||
2944 -----ACTAATGAATGATTTAGGAAGGCAAGAAAGACAGTGAATGATATTT 2994
729 ThrAsnHisThrIleThrLeu 735
|||:|||||
2995 ATTAACGCAACAAACGCTTTG 3015

RESULT 7
US-10-929-754-2
; Sequence 2, Application US/10929754
; Publication No. US20050124603A1
; GENERAL INFORMATION:
; APPLICANT: ABDULLAH, MODH AMIR
; APPLICANT: DEAN, DONALD H.
; TITLE OF INVENTION: INSECTICIDAL CRYSTAL PROTEINS WITH ENHANCED TOXICITY
; FILE REFERENCE: 22727/04179
; CURRENT APPLICATION NUMBER: US/10/929,754
; PRIOR FILING DATE: 2004-08-30
; PRIOR APPLICATION NUMBER: 60/498,826
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: Patent In Ver. 3.2
; SEQ ID NO 2
; LENGTH: 3684
; TYPE: DNA
; ORGANISM: Bacillus thuringiensis
US-10-929-754-2
Alignment Scores:
Pred. No.: 4,936-70 Length: 3684
Score: 759.00 Matches: 234
Percent Similarity: 47.4% Conservative: 122
Best Local Similarity: 31.2% Mismatches: 267
Query Match: 19.6% Indels: 128
Gaps: 34
DB:
US-10-783-417-2 (1-735) x US-10-929-754-2 (1-3684)
QY 23 AsnArgAsnSerAsnAspSerArgTyrProTyrThrAsnAsnProAsnGlnProLeuGln 42
|||:|||||
DB 154 AATATGAATTCAGC-----TATCCGTTAGCGAATGACTTCAAGGCTCAATGA 204
QY 43 AsnThrAsnTyrIleGluTyrLeuAsnMetCysGlnGlyAsnThrGlnTyrGlyAspAsn 62
|||:|||||
DB 205 AACGAACTATTAAGATTTGCTAGCCATGCTGTGAATTAACCAACAGATAGCTTAAT 264
QY 63 PheGluThrPheAlaSerAlaAspThrIleAlaValIleSerAlaGlyThrIleValSer 82
|||:|||||
DB 265 -----CCAGCTGCATTAATTTCTTCAATTAAGTACCGCTTAAAGTACT 312
QY 83 GlyThrLeuLeuAlaGlyIleGlyIleThrIleSerIleSerGlyProIleGlyIleIle 102
|||:|||||
DB 313 GAGAGTATC-----CTTAATTTGTAACCAACCTGCAGTACTGTC 354
QY 103 GlyAlaIleIleIleSerPheGlyThrLeuIleThrValPheThrProAlaGly---Glu 121
|||:|||||
DB 355 TTAACCTGACTT-----AGCGGAGTCTTCTTATTTGCGCCAGCTAATACTCA 405
QY 122 GlnAspIleThrValTyrThrGlnPheIleIleIleIleIleIleIleIleIleIleIle 141
|||:|||||
DB 406 ACGCTGAAGAAGTTTGAAATGATTTCAATGACCAATACAGGAACTTAATGATCAAACT 465
QY 142 LeuThrGluSerIleIleGlyGlnLeuIleIleIleIleIleIleIleIleIleIleIle 161
|||:|||||
DB 466 GTAACAGCTTATGACGAACAGATGCAAAATGCAAAATGACCGTTGTGAATATTTTA 525
QY 162 GlnSerTyrAsnThrAlaLeuAspAspTyrPargIleIleIleIleIleIleIleIleIle 181
|||:|||||
DB 526 GATCAATATACCACTTAATTTTAACCTTGAAAAA----- 561
QY 182 LeuProProSerSerAlaLeuGlnGlnAlaIleIleIleIleIleIleIleIleIleIle 201
|||:|||||
DB 562 GAGCTTAATTAACAGTCTTATGAAACAGACAGATTAATCT-----CAATTT---AAC 609
QY 202 ValHisAsnAspPheIleArgGluIleProGlyPhe-----GlnLeuGluThrTyrIle 219
|||:|||||
DB 610 TTAACCAAGTGCMAACTTCGAGAGACCGCAGTTTATTTTACCACTTATGATGATGA 669
QY 220 ThrLeuLeuLeuProIleTyrAlaGlnAlaIleAsnPheIleIleIleIleIleIleIle 239
|||:|||||


```

Db      670 TTATTGTTATTACCAATATACGACAGTACGCAAAATTCATTATTAATAAGAAAT 729
Qy      240 GYALAGLUEVALAASPGLUTRPAENLAEPILIEHISPROSEGLINILEGLUPROASN 259
Db      730 GGGCCATTAATAATGACACAGAAATGCTCT-----TTACACAGTACT 768
Qy      260 AAGLYTHSERASPSPTYRTRYLYSEULEULYSGLUASNILEPROLYSTYSEASN 279
Db      769 GCTGCT-----GACCAACTATATACTATGATGACAGTACATAAGAAATATATGCA 822
Qy      280 TYRCYSAIAENHTRTYRARGHTRGLYLEULYSAENLEUARGASGLUPROASMETYLS 299
Db      823 CATACACTTACATGGGTATATAAAGGTTTAGTACTTAGAAATAATCTAATAGAACAA 882
Qy      300 TTPSERILEPHEASAPRYARGARGTYRMECTHRIETHVALILEUASPTHRISESER 319
Db      883 TGGATTACGTTTAAAGATTATTAATAAGAGATGACTATTCAAGTATTAATTAATCTCGCT 942
Qy      320 GINPHESELEUTYRASPILLEYSAARGTYR---ARGASPSERILEGLYGLYLEGLUVAL 338
Db      943 CTTTGTGCGAGTTATGATCAGTCGATACCTGCGGCAAAATAGATATACGAAACTA 1002
Qy      339 LYSGLYLELYEASNGLUENHTRARGULIETRYRTHTRHGLUULEASNPHEASARG 358
Db      1003 TCA-----AAAACAGAAATTTACAGAGATTTATACAGCT----- 1038
Qy      359 LEUPROGLINLEUARGVALGINPROASNLEUALATHMETGLUTRYASMLEUTHTRARGALA 378
Db      1039 TTAGTAGAATCTCCTTCTAGTAATCTATAGCAGCACTGAGCAGCACTTACACAGA--- 1095
Qy      379 SERPHELYLEUPHESERHPELEUGLINPHEILEPHETYRTHGLUASNPHEASNPHE 398
Db      1096 GATGTTCAATTATTCACCTTGCTTAAGAGAGATGATTTCTGAGCC---AATACTATATAT 1152
Qy      399 GLYASNAAGLUEVALGLYLESERASNPARGAPALAPROTHTRYSERASNPHTHRIETHR 418
Db      1153 CAAGATTTAAGATTTTATATCTGCAATATAAATGGGTTTCAATACAAATCTTCTGCA 1212
Qy      419 -----GLUTHTHLEUTRYGLULARG---THRGLYSERPROTHTRYTHRYTHRIE 434
Db      1213 ATGCAGAAAGATGCAATTTATGGAAGTTCTGCTTTGGTTCAAACTTACTATCAAAATT 1272
Qy      435 ARGPROPEGLUSER---TYRILEVALSERILEVALTHIRASARGLINSEPROBROVAL 453
Db      1273 CAACTTAATTCTAATGTTATATAAATTCTATC---ACAGATACAGCTCCCC----- 1323
Qy      454 SERPROILEGINPROHISPEHILEILEANGINILEGLULEUTRYLEUASNGLYSER 473
Db      1324 -----TCT 1326
Qy      474 AENAHNTHLEULYS-----TYRSEALAGLYSERLEUSERASNPHTYRGINAN 490
Db      1327 AATCGAGTTACAAAATGATTTCTACAAAATGATGATGCTTGGCTTATATATCA 1386
Qy      491 -----THTRPHEPHEGLINPHEPROARGLYSELYS 500
Db      1387 AATATAACACCAACTCTGAAGGTTTAAGACCAACATTTTGGATTTTCAACAATAGAG 1446
Qy      501 AEPCYAENLEUVALILEASPPROGLYCYSERPROASNPHEASNPHTYRSEHISILE 520
Db      1447 AAC-----ACACCTAAT---CAACCAACTGTAAAGATTAATATCGCATATTT 1488
Qy      521 LEUSERHISPHESERLEUPHERHTRYSETRYVALILEGLYLEUGLINLEUINLEU 540
Db      1489 TTAACCTAT-----ATMAAACTGATGTATTA 1515
Qy      541 AEPTRGLY-----VALLEUGLYTRPHTHISERSESEVALASPARGYRAN 556
Db      1516 GATTATAACAGTAACAGGGTTTCATTGCTTGACACATAGATGTTGACCCCTAATAT 1575
Qy      557 AIAIISERASPLYILEIETHMETILEPROAIAILELYEGLYASNAENLEUASPTH 576
Db      1576 CAATATACACAGATGCTATCACACAGTTCCGGCGTAAATCTAATCTTGAATGCA 1635

```

```

Qy      577 AENSERLYSVALILEGLULYPROGLYHISTHTRGLYGLYASNLEUVALTRYLEUGLINSE 596
Db      1636 ACAGCTAAAGTATATCAAGACAGCTGTGCTATACAGGGGGGAGTCTAGTCTTACAAAC 1695
Qy      597 -----GINGLYARGLEUGLINLEUTHRYCYSGIUTHPRO-----AENSER 609
Db      1696 AATGCTACTCTATACAGCAGACATGAGATGATCAATGTAAACAAAGTATTTATGATCCT 1755
Qy      610 THRGINSETRYRPHLEIARGLEUARGYRPLATHTRANGLYALAGLYASNPHTHLEUPRO 629
Db      1756 ACAAGAACTTACGGATTACGATACGATGCTGCAAAATGATCCA-----ATTGTATG 1809
Qy      630 AENILESELEUTHRIEPROGLYVALILEGLYLEPROGLINARGLEUASNPHTH 649
Db      1810 AATGATATATATGATTAATTAACAAGAGTTCTCTAGAGAACACAGATTAGTACGAATCTACG 1869
Qy      650 PHESERGLYTHRAENTYR-----ASNAENLEUGINTRYGLYASPPHEGLYTRYRPE 666
Db      1870 TTTTCAAGACCTTAATATATATATACCTACACATTTAAATATAGAGAGTTTATGATACAA 1929
Qy      667 GIN---PHEPROSETHRYVALTHLEUPROLEUASNPARGAN-----ILEPROHEILE 683
Db      1930 GATCCTTTGATGCAATTTGATCCAGATGAGATTATCTTATACAACTGATTAATCTATACCT 1989
Qy      684 PHEASNPARGALASPPVAL---SERASNPHTHLEUULEIILEASPLYILEGLUPHEILE 702
Db      1990 ATTCAACCATTAATACATGACTTCAATATATCAAGATATTATGACAGATCAAAATTAAT 2049
Qy      703 PROILETHSERSEMECHIEGLASNPARGULYSGINLYSELEUGLUTHRIEGLINTHR 722
Db      2050 CCAATCACTCAATCTGATTTAGATGATGACAGCAAGAAACCAAAATTTAGATCAGAACGAGAA 2109
Qy      723 LYSILEAHNTHRPHETHPHETHASNPHTHRYLS 733
Db      2110 GTTGAAATGCACTGTTTACAAATGACGCGAAA 2142

RESULT 8
US-10-929-754-4
; Sequence 4, Application US/10929754
; Publication No. US20050124803A1
; GENERAL INFORMATION:
; APPLICANT: ABDULAH, MODH AMIR
; TITLE OF INVENTION: INSECTICIDAL CRYSTAL PROTEINS WITH ENHANCED TOXICITY
; FILE REFERENCE: 22727/04179
; CURRENT APPLICATION NUMBER: US/10/929,754
; PRIOR FILING DATE: 2004-08-30
; PRIOR APPLICATION NUMBER: 60/498,826
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: Patent In Ver. 3.2
; SEQ ID NO 4
; LENGTH: 4391
; TYPE: DNA
; ORGANISM: Bacillus thuringiensis
US-10-929-754-4

Alignment Scores:
Pred. No.: 4,48e-64 Length: 4391
Score: 705.00 Matches: 225
Percent Similarity: 43.4% Conserved: 119
Best Local Similarity: 28.4% Mismatches: 278
Query Match: 18.2% Indels: 171
DB: 9 Gaps: 29

US-10-783-417-2 (1-735) x US-10-929-754-4 (1-4391)
Qy      1 METLENGINMAENASNPASNPASNGLYTRYGLUIELLEASPSERHISTHRYSEPROTYR 20
Db      719 ATGCATATATTAATGCGAAATGAGAAATGAATATGACATATTA----- 757
Qy      21 PHEPROASNPARGASNPASNPASER-----ARGTRYPROTYR 33

```

Db 758 -----AATGCTTCATCAACGATTCAACATGTCTAATACTTATCCGAGTATCCGTTA 811
 QY 34 ThrAsnAsnProAsnGlnProLeuGlnAsnThrAsnTyrlsGluThrLeuAsnMetCys 53
 Db 812 GCAAATCCACCAACAGATTAAATGCAAAATTAACAAATTAATGAAGATTGGCTTAATGTATGT 871
 QY 54 GlnGlyAsnThrGlnTyrlsAsnPheGlnThrPheAlaSerAlaAspThrIleAla 73
 Db 872 GAAGGCTAT-----CATATAGAAAATCCATAGAAACCAAGCGTTAGAGCT 916
 QY 74 AlaValSerAlaGlyThrIleValSerGlyThrLeuAlaGlyIleGlyGlyLeuThr 93
 Db 917 GCACCTGGTAAGGAGATTAGCTATAGTATAGACATGAGGTGCTTGTGTGTGT--TCT 973
 QY 94 SerIleSerGlyProIleGlyIleIleGlyAlaIleIleIleSerPheGlyThrLeuIle 113
 Db 974 AATTATCTTAAGATACAAATGGATTG-----TTTATCCAAATATTCA 1012
 QY 114 ThrValPheThrProAlaGlyGluGlnAspLysThrValThrThrGlnPheIleLysMet 133
 Db 1013 GAGCTACTTGGCCAGAGATGATNACCCAGCAATACATTGGCAAGATATTATGAATCAT 1072
 QY 134 GlyGluIlePheValAspThrProLeuThrGlnSerIleLysGlnLeuLysLeuGlnThr 153
 Db 1073 GTAGAAGATCTTATAGACAAAGAAATTAAGTACGATTATACGAGAAATCCATTTAGACA 1132
 QY 154 LeuGluGlyPheArgGlnIleLeuGlnSerTyrlsAsnThrAlaLeuAspAspThrArgLys 173
 Db 1133 TTAGCAGATTATACAGGGTAAAGTGTATGATTATTAACAATTGGTTAGAAATGGAAAGAC 1192
 QY 174 LeuLysArgLeuGlnAlaIleProGlyLeuProProSerSerAlaLeuGlnIleAlaLeu 193
 Db 1193 -----GATCCAAATCTACAGGTAAATTAAAGC----- 1219
 QY 194 ThrLeuLysIleArgPheGlnAsnValIleAsnAspPhe-----IleArgGluIle 210
 Db 1220 ACCTTAGTACCAAGATTAGCGCTCTGTATTCAGATTTTAATGTCCTTAAGACAGATT 1279
 QY 211 -----ProGlyPheGlnLeuGlnGluThrTyrlsThrLeuLeuLeuProIle 225
 Db 1280 AATAATACAGGGAGATCCAGGTATGAGTTA-----CTTTATTTGGCTGTGC 1324
 QY 226 TyrAlaGlnAlaIleAsnPheIleAsnLeuLeuGlnGlnGlyAlaGluLeuAlaAsp 245
 Db 1335 TATGCACAAAATTCGCAATCTGATCTTATTAATTAAGGATGCTCAGATTTATGAGAGAT 1384
 QY 246 GlnUTPAsnAlaAspIleHisProSerGlnIleGluProAsnAlaGlyThrSerAspAsp 265
 Db 1385 AATATG-----TGAGCGCACGAGCTAATCTCTGTATAT 1420
 QY 266 TyrTyrlsLeuLeuLysGlnAsnIleProLysTyrlsSerAsnTyrlsCysAlaAsnThrTyrls 285
 Db 1421 TATTCACCAATATCAATTAGAAAACCAAGAAATATACAGAAATATGTTATTAATGGTAT 1480
 QY 286 ArgThrGlyLeuLysAsnLeuArgAspGluProAsnMetLysThrSerIlePheAsnAsp 305
 Db 1481 AATAAGGCTTTAAATGATTTTGAACACAGAGT-----CAATGGGTAAACCTTATATCT 1534
 QY 306 TyrArgArgTyrlsMetThrIleThrValIleAspThrIleSerGlnPheSerLeuTyrlsAsp 325
 Db 1535 TATCTGTAAGAAATGATCTTACTGTATTAAGTATTAATTAATTAATTAATTAATTAATTAAT 1594
 QY 326 IleLysArgTyrlsArgAspSerIleGlyGlyIleGluValIleGlyIleLysAsnGlnLeu 345
 Db 1595 GCGAGATTATATCTTACAGAA-----GTAAAAACCGAATCA 1630
 QY 346 ThrArgGluIleTyrlsThrThrGluIleAsnPheAspArgLeuProGlnLeuArgValGln 365
 Db 1631 ACTTAGGAAATTTATTCAGATGTTATTAATGGGAGATATATGACATTATGACTCTTAT 1690
 QY 366 ProAsnLeuAlaThrMetGluTyrlsAsnLeuThrArgAlaSerPheLysLeuPheSerPhe 385

Db 1691 TTTCTTTTGAGAAAGCTGAATCACTTATATCAAGGACCC-----CATCTTTCACCTTGG 1747
 QY 366 LeuGlnGlnPheIlePheTyrlsGluAsnThrAsnPheGlyAsnArgLeuValGlyIle 405
 Db 1748 CTAAAGATTTTCATTTGTATGACCAATTCATTTCTTATGACATTTTATACAGGTGT 1807
 QY 406 SerAsnArgAspAlaProThrTyrlsSerAsnThrIleThrGlu----- 419
 Db 1808 CAATATAGATTTCTTATACTAATATCTAGATTTAAGAGCGCTTTTATGGGAGACG 1867
 QY 420 ---ThrLeuTyrlsGluArgThrGlySerProThrThrLysThrIleArgProPheGlu 438
 Db 1868 GACACAGATTAATGT-----GGGACTTCTTACCATTAATTAATTT--CCATCAAT 1915
 QY 439 SerTyrlsValSerIleValThrAspArgLys-----SerProPro 452
 Db 1916 TCGTATGTATATTAATTAATGACGAAATTAATGAATTAATTAATTAATTAATTAATTAATTAAT 1975
 QY 453 ValSerProIleGlnProHisPheIleIleAsnGlnIleGluLeuTyrlsLeuAsnGlySer 472
 Db 1976 GTAAATATTTACAAAATGAATTTTCTGTACA-----GATAATAT 2017
 QY 473 SerAsnAsnThrLeuLysTyrlsSerAlaGlySerLeuSerAsnTyrlsAsnThrThr 492
 Db 2018 TCTTCAAAAGATTAATTAATTAATGTCACACAGAAAGAAATTAACCTGTTGTCGACAGAT 2077
 QY 493 PhePheGlnPheProArgLysLysAspCysAsnLeuValIleAspProGlyCysSerPro 512
 Db 2078 TTTGATTTTCTCACATTAATGAAGGAACTGAGTTAGCA----- 2116
 QY 513 AsnPheAsnAsnTyrlsSerHisIleLeuSerHisPheSerLeu----- 526
 Db 2117 AATATTAATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2176
 QY 527 -----PheThrTyrlsSerTyrlsValIleGlyLeuGlnLeuGlnIleLeuAspThrGly 543
 Db 2177 CAGAAACGTCAATGTTATTCGTT----- 2200
 QY 544 ValLeuGlyThrPheHisSerSerValAspArgTyrlsAsnAlaIleSerAspLysIleIle 563
 Db 2201 -----GCTTTTACATAGTAGTGTTGATCTTAATTAATCACTTACGCAATTAAT 2254
 QY 564 ThrMetIleProAlaIleLysGlyAsnAsnLeuAspThrAsnSerLysValIleGlyGly 583
 Db 2255 ACGCAATTTCTGTATGTAAGCTTCAGATTAATTAATGATCATTTCAATTAATAAAGGT 2314
 QY 584 ProGlyHisThrGlyGlyAsnLeuValTyrlsLeuGlnSerGlnGlyArgLeuGluIleThr 603
 Db 2315 CCCGATTTTACGGGAGAGATTTGGTAAAGATGAAGCAGATTACGTTTAATCTATGGT 2374
 QY 604 CysGluThrProAsnSerThrGlnSerTyrlsPheIleArgLeuArgTyrlsAlaThrAsn--- 622
 Db 2375 TTTAAGCTGAATTTTATGATTAATAAATATCTGTTCAATATGTTAATATGTAATCA 2434
 QY 623 -----GlyAlaGly-----AsnThrLeuPro 629
 Db 2435 AGTTCTAAATTAATTAATCAAGAAATGAAGGAGGATTAATTAATCAACAAATTAATCA 2494
 QY 630 AsnIleSerLeuThrIleProGlyValIleGlyIleProProGlnArgLeuAsnAsnThr 649
 Db 2495 AATATTTCTCCACA----- 2509
 QY 650 PheSerGlyThrAsnTyrlsAsnLeuGlnTyrlsGlyAspPheGlyTyrlsPheGln---Phe 668
 Db 2510 -----TATGAGCGCTTTTCTTATTAATGAGTCTTT 2539
 QY 669 ProSerThrValThr-----LeuProLeuAsnArgAsnIleProPheIlePheAsn 685
 Db 2540 ACTATTAATCAAGACAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2599
 QY 686 ArgAlaAspValSerAsn---SerIleLeuIleIleAspLysIleGluPheIlePro-I1 704
 Db 2600 AGACAGTTTGTGAAGATATACATCTCTTATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2659

```
Qy 704 eThrSerSerMetHsGlnAsn--ArgGlnLysGlnLysLeuGlnThrIleGlnThrly 723
Db 2660 AACTGATACCATTCACAGGAATATGAGGAAATATGAAATATGAAAGATCACAGGAAC 2719
Qy 723 sTleAsnThrPhePheThrAsnHisThrlyThrleu 735
Db 2720 ATTATATGATATTTGTTGATTAAACAAAGTACTA 2756

RESULT 9
US-10-032-717-3
; Sequence 3, Application US/10032717
; Publication No. US20020151709A1
; GENERAL INFORMATION:
; APPLICANT: Andre R. Abad
; APPLICANT: Nicholas B. Duck
; APPLICANT: Xiang Feng
; APPLICANT: Ronald D. Flannagan
; APPLICANT: Theodore W. Kahn
; APPLICANT: Lynn E. Sims
; TITLE OF INVENTION: Genes Encoding No. US20020151709A1e1 Proteins With
; TITLE OF INVENTION: Pesticidal Activity Against Coleopterans
; FILE REFERENCE: 35718/237005
; CURRENT FILING DATE: 2001-10-23
; PRIOR APPLICATION NUMBER: 60/242,838
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 3633
; TYPE: DNA
; ORGANISM: Bacillus thuringiensis
; FEATURES:
; NAME/KEY: CDS
; LOCATION: (1)..(3633)
; NAME/KEY: misc_feature
; LOCATION: (0)..(0)
; OTHER INFORMATION: CRY1218-2
US-10-032-717-3

Alignment Scores:
Pred. No.: 5.34e-64 Length: 3633
Score: 703.00 Matches: 219
Percent Similarity: 44.5% Conservative: 127
Best Local Similarity: 28.2% Mismatches: 289
Query Match: 18.2% Indels: 142
Db: 5 Gaps: 32

US-10-783-417-2 (1-735) x US-10-032-717-3 (1-3633)
Qy 5 AsnAspAsnAsnGlnGlnIleIleAspSerHisThrSerProTyrrPheProAsnArg 24
Db 10 AATATATCAAAATGAAATATGAAATATATAGATGCGACACTTCT-----ACTTCT 57
Qy 25 AsnSerAsnAspSer---ArgTyrrProTyrrThrAsnAsnProAsnGlnProLeuGlnAsn 43
Db 58 GATACCAATGATTTCAACAGATACCTTTTGCGAATGAGCAAAATGCGCTACAAAT 117
Qy 44 ThrAsnTyrrLysGlnThrLeuAsnMetCysGlnGlnAsn---ThGlnTyrrGlnAsn 62
Db 118 ATGGAATTATAAAGATTATTTAAATGCTCGCGGAAATGCTAGTAATACCTCGTTCA 177
Qy 63 PheGlnThrPheAlaSerAlaPheThrIleAlaIleAlaValSerAlaGlyThrIleValSer 82
Db 178 CCTGAAGTACTGTTTACCGGACACAGAT-----GACGCTAAGGCGCAATTGATATGTA 231
Qy 83 GlyThrLeuLeuAlaGlyIleGlyLeuThrSerIleSerGlyProIleGlyIleIle 102
Db 232 GGTAAATTACTATCAGGTTTAGCG-----GTCCATTGTT 267
Qy 103 GlyAlaIleIleIleSerPheGlyThrLeuIleThrValPheTyrProAlaGlyGlnGln 122
```

```
Db 268 GGGCCGATAGTAGTCTTTATATCTCACTTATGTGATTTCTGTGCGCTTCAAGGCGAA--- 324
Qy 123 AspLysThrValTyrThrGlnPheIleLysMetGlyGlnIlePheValAspThrProLeu 142
Db 325 ---AAGACTCAATGGGCAATTTTATGGAACAAGTGAAGAACTGATTAATCAAAAAATA 381
Qy 143 ThrGlnSerIleLysGlnLeuLysLeuGlnThrLeuGlnGlyPheArgGlnIleLeuGln 162
Db 382 GCAGAAATATGCAAGGAATTAAGCCCTTCGGAATTAGAAGGATTAGTAAATTAATACCA 441
Qy 163 SerTyrrAsnThrAlaLeuAspAspThrPheArgLysLeuLysPheArgLeuAlaProLysLeu 182
Db 442 TTATATCTTAACTGCGCTTGAAGAAATGAAGAA----- 474
Qy 183 ProPheSerSerAlaLeuGlnGlnAlaIleLeuThrLeuLysIleArgPheGlnAsnVal 202
Db 475 AATCCAAATGGTTCA-----AGAGCCTTAGCAAGATGTCGAAATCGAATTTGAAATCTCG 528
Qy 203 HisAsnAspPheIleArgGlnIleProGlyPheGlnLeuGlnThrTyrrLysThrLeuLeu 222
Db 529 GATAGTTTATTTACCGCAATATACATGCGATCTTTTCGAGTGAACAAATTTGAAAGTACCATTC 588
Qy 223 LeuProIleTyrrAlaGlnAlaIleAsnPheHisLeuAsnLeuLeuGlnGlnGlyAlaGln 242
Db 589 CTTACAGTATATACACAGGACGCAACCTTCAATTTACTGTTATTAAGAGACGCTTCAAT 648
Qy 243 LeuAlaAspGlnTyrPheAsnAlaAspIleHisProSerGlnIleGlnProAsnAlaGlyThr 262
Db 649 TTGAGAGAAAGATGGGATGG-----TCTAACACCACT 681
Qy 263 SerAspAspTyrrTyrrLysLeuLeuLysGlnAsnIleProLysTyrrSerAsnTyrrCysAla 282
Db 682 ATTATATACCTATTTATGATCGTCAATGAACCTTATGCAAGATATTCGATCATCGTGTGA 741
Qy 283 AsnThrTyrrArgThrGlyLeuLysAsnLeuArgAspGlnProAsnMetLysThrSerIle 302
Db 742 AAGGTATGAAACCTGTTTATGCAAAATTTAAAGGACAGACGCGCTAAACAGGCTCAAC 801
Qy 303 PheAsnAspTyrrArgArgTyrrMetThrIleThrValLeuAspThrIleSerGlnPheSer 322
Db 802 TATTAACCAATTCGAGAGAAATGACATGACGCTTTAGATGTTGTTGCAATATATCCCA 861
Qy 323 LeuTyrrAspIleLysArgTyrrArgAspSerIleGlyGlyIleGlnValLysGlyIleLys 342
Db 862 AATTATGACACACGACGCTAC-----CCAATGAAACCAAGACA----- 900
Qy 343 AsnGlnLeuThrArgGlnIleTyrrThrThrGlnIleAsnPheAspArgLeuProGlnLeu 362
Db 901 ---CAACTAACAGGAAGATATATCA-----GATCCACTGGGCGCGGTGA 942
Qy 363 ArgValGln-----ProAsnLeuAlaThrMetGlnTyrr 373
Db 943 AACTGTCTTCAATGCTTCCGCTGATACAAAGACCTTCTTTCGGAGCTATATACATCA 1002
Qy 374 AsnLeuThrArgAlaSerPheLysLeuPheSerPheLeuGlnGlnPheIlePheTyrrThr 393
Db 1003 TCCGTTATTCAGACCAACC---CATGTATTTGATTATATACGCGGACCTACAGTATATCA 1059
Qy 394 GluAsnThrAsnPheGlyAsn-----ArgLeuValGlyIleSerAsnArg 408
Db 1060 CAATCAAGACGATTTCTTCGCTCGCTATATTAACATTTGGCGCTGCTATCAATTAAGC 1119
Qy 409 AspAlaProThrTyrrSerAsnThrIleThrGlnThrLeuTyrrGlyLysArgThrLysSer 428
Db 1120 TATCATCGGATTTTATGATATATTTATTAACAGATGATATGAACTATCAAAATCTA 1179
Qy 429 ProThrThrLysThrIleArgProPheGlnSerTyrr-----Lys 441
Db 1180 CACAGCACTAGTACTTT---GATTTACGAAATATGATATTTACAGACGTTATCAAAA 1236
Qy 442 ValSerIleValThrAspArgLysSerProProValSerProIleGlnProHisPheIle 461
Db 1237 GATGCGGTCCTTGAATATGTTTCTGTTATACGTATATA-----TTTTTTGA 1290
```

Oy		462	IleAsnGlnIleGluLeuThrLeuAsnGlySerSerAsnAn-----ThrLeuLys	478
Db		1291	ATGCCAGAAAGTCGAAGTCTTTTCATGTATAACCAATTGAATTAATACCAGAAAAGCGTTAAG	1350
Oy		479	TyrSer-----AlaGlyGlySerLeuSerAsnTyrglnAsnthr	491
Db		1351	TATATCCGGTTCCCAAAGATVATTATATACCGGGG-----	1383
Oy		492	ThrPhePheGlnPheProArgLysAlaPheCysAsnLeuValIleAspProGlyCysSer	511
Db		1384	-----ACAAGAATTCGGAATTAAGATTAACCTCCAGAAACTTCA	1422
Oy		512	-----ProAsnPheAsnAsnTyrsenHisIleLeuSerHisPhe-----SerLeuPheThr	528
Db		1423	GATCACCAACCAATTATAGACTCATATAGCCATATGATTAATGCATATACACAAAGATATCCC	1482
Oy		529	TyrSerTyrglnValIleGlyLeuGlnLeuGlnIleLeuAspThrGlyValLeuGlyTyrrThr	548
Db		1483	ACGGGTCACTACCGGATTA-----GTAACGTATTTCTTGAGCA	1524
Oy		549	HisSerSerValAspArgTyrglnAlaIleSerAspLysIleIleThrMetIleProAla	568
Db		1525	CATCGAATGCCGAATCTTAAATGACAGTTCATTCAGATAAAATTAATCAAGATTCGCGGTC	1584
Oy		569	IleLysGlyAsnAsnLeu-----AspThrAsnSerLysValIleGlu	582
Db		1585	GTAAGGTCTTCGATTTGGCTCCCTCTTAACAGAGAGGCCAAATTAATACCGTGTATCG	1644
Oy		583	GlyProGlyHisThrGlyGlyAsnLeuValTyrglnGlnSerGlnGly-----	598
Db		1645	GGTCCGTGATTTAACAGGGGGGGGAGTAAATAAAGAAATGAGATTAATATATCA	1704
Oy		599	-----ArgLeuGlnIleThrCysGluThrProAsnSerThrGlnSerTyrrPheIleArg	616
Db		1705	CATATCCGTGTTAAAAATTCA-----GACATTAACAAGATTAATATGATGAGG	1752
Oy		617	LeuArgTyrglnAlaThrAsnGlyValaGlyAsnThrLeuProAsnIleSerLeuThrIlePro	636
Db		1753	ATTCCGATTCGCTTC-----GCTAATATTAATCGAATTTTATATTAATCCTTCGAAACA	1806
Oy		637	GlyValIleGlyIleProGlnArgLeuAsnAsnthrPheSerGlyTyrrAsnTyrrAsn	656
Db		1807	AACGTTAATCT---CACGCTCAAAAACTATGATATAGAGGTGAAGCTTTAATCATATAT	1863
Oy		657	AsnLeuGlnTyrglyAspPheGlyTyrrPheGlnPheProSerThr-----ValThr	673
Db		1864	AAATTTAATTATGCGACTTTCGCCCTTATTAATTAATTAAGACAACCGAACCCTTCATTA	1923
Oy		674	LeuProLeuAsnArgAsnIleProPheIlePheAsnArgAlaAspValSerAsnSerile	693
Db		1924	CTTAGGGGCT-----ATATTGAAGCGGAAGACTTCTTGGAATTGAA	1965
Oy		694	LeuIleIleAspLysIleGlnPheIleProIleThrSerSerMetHisglnAsnArgGlu	713
Db		1966	GCTTAATATAGACCGAATCGAATTTATCCAGAGATGAGACATAT-----GAAGCG	2016
Oy		714	LysGlnLysLeuGlnUthrIleGlnThrLysIleAsnthrPhePheThrAsn	730
Db		2017	GAACAGATTATTAAGAGACGACGAAAGAACGATGAATGCTTTGATTAAGAAAT	2067
 RESULT 10 US-10-414-637-3 ; Sequence 3, Application US/10414637 ; Publication No. US20030177528A1 ; GENERAL INFORMATION: ; APPLICANT: Andre R. Abad ; APPLICANT: Nicholas B. Duck ; APPLICANT: Xiang Feng ; APPLICANT: Ronald D. Flannagan ; APPLICANT: Theodore W. Kahn ; APPLICANT: Lynn E. Sims ; TITLE OF INVENTION: Genes Encoding No. US20030177528A1e1 Proteins With				

```

1  TITLE OF INVENTION: Pesticidal Activity Against Coleoptera
2  FILE REFERENCE: 35718/237005
3  CURRENT APPLICATION NUMBER: US/10/414,637
4  CURRENT FILING DATE: 2003-04-16
5  PRIOR APPLICATION NUMBER: US/10/032,717
6  PRIOR FILING DATE: 2001-10-23
7  PRIOR APPLICATION NUMBER: 60/242,838
8  PRIOR FILING DATE: 2000-10-24
9  NUMBER OF SEQ ID NOS: 48
10 SOFTWARE: FastSeq for Windows Version 4.0
11 SEQ ID NO 3
12     LENGTH: 3633
13     TYPE: DNA
14     ORGANISM: Bacillus thuringiensis
15     FEATURE:
16     NAME/KEY: CDS
17     LOCATION: (1)...(3633)
18     FEATURE:
19     NAME/KEY: misc feature
20     LOCATION: (0)...(0)
21     OTHER INFORMATION: Cyt1218-2
22     US-10-414-637-3

```

Alignment Scores:

Pred. No.:	5,34e-64	Length:	3633
Score:	703.00	Matches:	219
Percent Similarity:	44.5%	Conservative:	127
Best Local Similarity:	28.2%	Mismatches:	289
Query Match:	18.2%	Indels:	142
DB:	6	Gaps:	32

US-10-783-417-2 (1-735) X US-10-414-637-3 (1-3633)

Oy		5	AAspAAspaAnagluyrGluIleLelelAsSerhIsthSrProtyrPheProAsnArg	24
Db		10	AAATAACCAAAATGAATATGAAAATTATAGTGGACACTTC-----ACTTCT	57
Oy		25	ASnSerASnAspSer---ArGlyrProTyThrkAnAsnProAngInProLeuGlnAsn	43
Db		58	GTAATCCAAATGATCTTTACAGATACCCTTTTTGGAAATGAGCACAAGAAATGGCGTACAAAT	117
Oy		44	ThrAsnTyrlYegluITrpleuAsnMeCyseGlnGlyAsn--ThrgIntyrGlyAspAsn	62
Db		118	ATGAAATTATAAGATTATTTAAAAATGTCTGGGGAATGTCTAGTAATACCTCGTTCTA	177
Oy		63	PheGIuThrPheAlasErLAspThrlleAlalavalIsErIaglyThrlleValSer	82
Db		178	CCTGAAGAATCTGTATACGGCACAGAT-----GCAGTAAAGGCCGCAATGATATATGTA	231
Oy		83	GLYThrLeuLeuAlaglylileGlyGlyLeuthrSerIleSerglyProIleGlylle	102
Db		232	GGTAATTTACTATCAGGTTTAGGG-----GTCCCATTTGTT	267
Oy		103	GlyAlaIleIleIleSerPheGlyThrLeuIleThrValPheTrpProAlaglyGluGln	122
Db		268	GGGCCGATAGTGAGCTTATATCACTCAACTATATGTATCTGTGGCCTTCAGGGCAA---	324
Oy		123	AspLystrValITrPhroInPheIlelySmetGlyIuleIephelvalAspThrProleu	142
Db		325	---MAAGATCAATGGGAGATTTTATYGACAACAGTGAAGAACATCTCAATATCAAAAATA	381
Oy		143	ThrgISerIleYegluIneulySreugInThrLeuGlnGlyPheArgGlnIleLeuGln	162
Db		382	GCAGAAATATGCAGAAATTAAGCCGCTTTCCGAAATTAGAGATTTAGTAATATACCA	441
Oy		163	SerTyAsnThrlAlaleuAspAspITrpArglySeulySarGleuGlnAlaProglyLeu	182
Db		442	TATATCTAACTCGCGCTTGAAGAAATGAAAGA-----	474
Oy		183	ProProSerSerAlaleuGlnGlnAlalAleuthrLeuysrIleArgPheGluAsnVal	202
Db		475	AATCCAATAAGGTTC-----AGACGCTTACGAGATGGCAAAATGCAATTTTGAATCTCG	528

QY 203 HisAsnAspPheIleArgGluIleProGlyPheGlnLeuGluThrTyrlsThrLeuLeu 222
Db 529 GATGATTTATTTACCAATACATGCCATCTTTCCAGTACCAAAATTTGATACCATTC 588
QY 223 LeuProIleTyraIaGlnIaIaAsnPheHlsleuAsnLeuGlnGlnIaIaGlu 242
Db 589 CTTACAGTATATACACAGGACGACCTTCATTTACGTATATTAAGAACCTTCAATT 648
QY 243 LeuIaAspGluTyrAsnIaAspIleHlsProSerGlnIleGluProAsnIaGlyThr 262
Db 649 TTTGGAGAAGATGGGATG-----TCTACAAACCACT 681
QY 263 SerAspAspTyrTyrlsLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 282
Db 682 ATTAATACCTATTAATGATGCTCAAAATTAAGAACTTACGATTTCTGATCCTGTGA 741
QY 283 AsnThrTyraGlyThrGlyLeuIaAsnLeuAraAspGluProAsnMetIysrIle 302
Db 742 AAGTGTATGAACCTGTGTTTACCAAAATTTAAAGCAGCGCTTAACATGGGTCCAC 801
QY 303 PheAsnAspTyraGlyTyraGlyThrIleThrIleValLeuAspThrIleSerGlnPheSer 322
Db 802 TATAACCATTCGATGAGAAATGACACGAGTTTATGATGTTGTCATATTCCCA 861
QY 323 LeuTyraPheIleIysArgTyraAspSerIleGlyIleGluValIleGlyIleIys 342
Db 862 AATTATGACACACAGCTAC-----CCAAATGAAACGAAACCA----- 900
QY 343 AsnGluLeuThrArgGluIleTyrlsThrIleGluIleAsnPheAspArgLeuProIleu 362
Db 901 ---CAACTACAAAGAGATATATACA-----GATCCACTGGGCGCGGTGA 942
QY 363 ArgValGln-----ProAsnLeuIaIaThrMetGlyTyrls 373
Db 943 AACGTCCTTCAATGTCCTCGTATAGCAAGACCTTCTTGGAGCTATTAATCA 1002
QY 374 AsnLeuThrArgAlaSerPheIleuPheSerPheLeuGlnPheIlePheTyrls 393
Db 1003 TCCGTTATTCGACCAACC---CATGATTGATTATATACCGGACCTCAAGTGTATCA 1059
QY 394 GluAsnThrAsnPheGlyAsn-----ArgLeuValGlyIleSerAsnArg 408
Db 1060 CAATCAAGAGCATTTCTTCGCTGCTATATTAAGACATTTGGCTGGTATCAAAATAC 1119
QY 409 AspAlaProThrTyraSerAsnThrIleThrGluThrIleGlyIleValArgThrGlySer 428
Db 1120 TATCATCGGATTTTATGATATATTAATAAAGATGATGAACCTTAATGAATCA 1179
QY 429 ProThrThrIysThrIleArgProPheGluSerTyrls-----Iys 441
Db 1180 CACACACACTAGTACCTTT---GATTTACGATTTATGATATTTACAGACGTTATCAAAA 1236
QY 442 ValSerIleValThrAspArgIleSerProProValSerProIleGlnPheIle 461
Db 1237 GATGGCGGTCTCTTGATATTTGTTTCTCGTATATACCTATATA-----TTTTTGGCA 1290
QY 462 IleAsnGlnIleGluLeuTyrlsAsnGlySerSerAsnAsn-----ThrLeuIys 478
Db 1291 ATGCCAAGAGTCGAGTTTTCATGTGAACCAATTGAATTAATCAAGAAAGCGTTAAAG 1350
QY 479 TyrlsSer-----AlaGlyIleSerLeuSerAsnTyrlsAsnThr 491
Db 1351 TATATCCGCTTCCAAAGATATATTAAGCGGG-----TCTACAAACCACT 1383
QY 492 ThrPhePheGlnPheProArgIysIysAspCyAsnLeuValIleAspProGlyCySer 511
Db 1384 -----ACAAGAGATTCCGAATTAAGAAATTACCTCCCAAAACCTTCA 1422
QY 512 -----ProAsnPheAsnAsnTyrlsSerHisIleLeuSerHisPhe-----SerLeuPheThr 528
Db 1423 GATCAACCAAAATTAAGATCATATAGCCATAGATTATGTCAATCAAGATATTTCCCGG 1482
QY 529 TyrlsSerTyrlsValIleGlyLeuGlnLeuGlnIleLeuAspThrGlyValIleGlyTyrls 548

Db 1483 ACGGTTCCAACTACCGGATTA-----GTAACCTGATATTTCTTGGACA 1524
QY 549 HisSerValAspArgTyraAsnIaIleSerAspIysIleIleThrMetIleProIa 568
Db 1525 CATCGAGTCCGATCTTATTAATGACGTTCACTTCAATTAATTAATCAATTCGAGTCC 1584
QY 569 IleIysGlyAsnAsnLeu-----AspThrAsnSerIysValIleGlu 582
Db 1585 GTAAGGTTCTGATTTGCTCCCTCTATTAACAGAGGCGCAAAATTAATACCGTTGATCG 1644
QY 583 GlyProGlyHisThrGlyIleAsnLeuValTyrlsLeuGlnSerGlnIle----- 598
Db 1645 GGTCTCGATTATTAACGAGGCGGAGTAAATAAGTAATAGAAATGAGTAATATATCA 1704
QY 599 -----ArgLeuGluIleThrCyseGluThrProAsnSerThrIleSerTyrlsPheIleArg 616
Db 1705 CATATGCTGTAAATTTCA-----GACATTAACAAAGATATATGATGAGG 1752
QY 617 LeuArgTyraIaIaIaAsnGlyIaGlyAsnThrIleuProAsnIleSerLeuThrIlePro 636
Db 1753 ATTCGATATGCTTC-----GCTAAATATCTGAATTTATATTAATCTTGAAGA 1806
QY 637 GlyValIleGlyIleProProGlnArgLeuAsnAsnThrPheSerGlyThrAsnTyrls 656
Db 1807 AACGTTAAATCT---CACGCTCAAAAACCTATGATAGGTGAACTTTAAATATAT 1863
QY 657 AsnLeuGlnTyraGlyAspPheGlyTyrlsPheGlnPheProSerThr-----ValThr 673
Db 1864 AAATTAATTAATGACGACTTGGCCCTTATTAATTAATTAACGACAAACCTTTCATCT 1923
QY 674 LeuProLeuAsnArgAsnIleProPheIlePheAsnArgAlaAspValSerAsnSerIle 693
Db 1924 CTAGCGGCT-----ATATTTGAAGCGAAACCTTCTTGAAATGA 1965
QY 694 LeuIleIleAspIysIleGlnPheIleProIleThrSerSerMetHisGlnAsnArgGlu 713
Db 1966 GCTTATATAGACGAGATGAATTAATCCAGTATGATGACATAT-----GAACGG 2016
QY 714 LysGlnIysLeuGluThrIleGlnThrIysIleAsnThrPhePheAsn 730
Db 2017 GAACAAATTTAGAAAGCAGCAAGAAAGCAATGATGCTTGTTCAGAAAT 2067

RESULT 11
US-10-606-320-3
; Sequence 3, Application US/10606320
; Publication No. US20040091505A1
; GENERAL INFORMATION:
; APPLICANT: Andre R. Abad
; APPLICANT: Ronald D. Flannagan
; APPLICANT: Rafael Herrmann
; APPLICANT: Theodore W. Kahn
; APPLICANT: Albert L. Lu
; APPLICANT: Billy Fred McCutchen
; APPLICANT: James K. Presnall
; APPLICANT: James F.H. Wong
; APPLICANT: Cao-Guo Yu
; TITLE OF INVENTION: Gene Encoding Proteins With Pesticidal
; TITLE OF INVENTION: Activity
; FILE REFERENCE: 35718/263948
; CURRENT APPLICATION NUMBER: US/10/606,320
; PRIOR FILING DATE: 2003-06-25
; PRIOR APPLICATION NUMBER: 60/391,766
; PRIOR FILING DATE: 2002-06-26
; PRIOR APPLICATION NUMBER: 60/460,787
; PRIOR FILING DATE: 2003-04-04
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO. 3
; LENGTH: 3633
; TYPE: DNA
; ORGANISM: Bacillus thuringiensis
; FEATURE:

Oy	599	-----AagLeuGluIleThrCysGluThrProAsnSerThrGlnSerTyrPheIleArg	616
Db	1705	CATATGCGTGTAAATAATTCA-----GACATTAAACAAAGATATAGTAGTAGG	1752
Oy	617	LeuArgTyrAlaThrAsnGlyValGlyAsnThrLeuProAsnIleSerLeuThrIlePro	636
Db	1753	ATTGCGATGCTTCC-----GCTAATATACTGAATTTATATATAATCCTTCGAAAGA	1806
Oy	637	GlyValIleGlyIleProProGlnArgLeuAsnAsnThrPheSerGlyThrAsnTyrAsn	656
Db	1807	AACGTTAATCT---CACGCTCAAAAACCTAGATAGAGGTGAAGCTTTAAACATATAT	1863
Oy	657	AsnLeuGlnTyrGlyAspPheGlyTyrThrGlnPheProSerThr-----ValThr	673
Db	1864	AAATTTAATTATGCGACTTGTCCCTTATTTAAATTTAGCAACCGAACTTTCACTTACT	1923
Oy	674	LeuProLeuAsnArgAsnIleProPheIlePheAsnArgAlaAspValSerAsnSerIle	693
Db	1924	CTAGGGGCT-----ATATTTGAAGCGGAAGACTTTCTTGGAATTGAA	1965
Oy	694	LeuIleIleAspTyrIleGlnPheIleProIleThrSerSerMetCHisGlnAsnArgGlu	713
Db	1966	GCTTATATAGACCGAATGCAATTTATCCAGTAGATGACATAT-----GAAGCG	2016
Oy	714	LyseGlnTyrLeuGlnThrIleGlnThrTyrIleAsnThrPheIleThrAsn	730
Db	2017	GAACAAAGATTTAGAAGCAGCAGCAAGAAACCAAGATGCTTGTTTACCAAT	2067

[illegible]

```

Db      1966 GCTTATATTAGCCGAATCGAATTATTCACGATGATGACAATAT-----GAAGCG 2016
OY      714 LysGlnLysLeuGIUThrIleGIThrLysIleAsnThrPhePheTrpAsn 730
       ::::|::::|::::|::::|::::|::::|::::|::::|::::|::::|
Db      2017 GAACAAGATTATTAAGAAGCAGCAAGAAAGAACGATGAAATCCTTGTATACAAAT 2067

RESULT 13
US-10-032-717-28
; Sequence 28, Application US/10032717
; Publication No. US20020151709A1
; GENERAL INFORMATION:
; APPLICANT: Andre R. Abad
; APPLICANT: Nicholas B. Duck
; APPLICANT: Xiang Feng
; APPLICANT: Ronald D. Flanagan
; APPLICANT: Theodore W. Kahn
; APPLICANT: Lynn E. Sims
; TITLE OF INVENTION: Pesticidal Activity Against Coleopterans
; TITLE OR INVENTION: Pesticidal Activity Against Coleopterans
; FILE REFERENCE: 35718/23705
; CURRENT APPLICATION NUMBER: US/10/032,717
; CURRENT FILING DATE: 2001-10-23
; PRIOR APPLICATION NUMBER: 60/242,838
; PRIOR FILING DATE: 2000-10-24
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 28
; LENGTH: 6613
; TYPE: DNA
; ORGANISM: Bacillus thuringiensis
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (0)..(0)
; OTHER INFORMATION: Genomic Cry12I8-2
US-10-032-717-28

Alignment Scores:
Pred. No.:          1,47e-63           Length:          6613
Score:              703.00             Matches:          219
Percent Similarity: 44.5%               Conservative:     127
Best Local Similarity: 28.2%            Mismatches:      289
Query Match:        18.2%                Indels:          142
DB:                  5                   Gaps:            32
US-10-783-417-2 (1-735) x US-10-032-717-28 (1-6613)

OY      5 AsnAspAsnAngIUThrGIUlleIleAspSerHisThrSerProTyrrPheProAsnArg 24
       |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db      1263 AATAATCAAATGGATATGAAATATATGATGCGCACACTTCT-----ACTTCT 1310
OY      25 AsnSerAsnAspSer---ArgTyrrProTyrrThrAsnAsnProAsnGlnProLeuGlnAsn 43
       |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db      1311 GTATTCGATGATGTTCTTACACAGTACCCTTTGCCAATGAGCCAACAATGCCTACAAAAT 1370
OY      44 ThrAsnTyrrIyrgIUThrPLeuAsnMetCysGlnGlyAsn--ThrgInTyrrGlyAspAsn 62
       ::::|::::|::::|::::|::::|::::|::::|::::|::::|::::|
Db      1371 ATGATATTATAAAGATTATTAAAAATGTCGCGGAATGCTAGTGAATACCCTGGTTCA 1430
OY      63 PheGIUThrPheAlaSerAlaAspThrIleAlaValaValaSerAlaGlyThrIleValSer 82
       |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db      1431 CCTGAAGTACTGTTTAGCGGAGCAAGAT-----GCACGTAAAGCGCGCAATTGATATAGTA 1484
OY      83 GIUThrLeuLeuAlaGlyIleGlyGlyLeuThrSerIleSerGlyProIleGlyIleIle 102
       |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db      1485 GGTAATATTACTATCATCGATTAGCG-----GTCCATTGTGT 1520
OY      103 GlyAlaIleIleIleSerPheGlyThrLeuIleThrValPheTrpProAlaGlyGluGln 122
       ::::|::::|::::|::::|::::|::::|::::|::::|::::|::::|
Db      1521 GGGCGGATAGAGACTTTTATACTCAACTTATATATATGTCGTGCGCCTTCAGGCGCA--- 1577
OY      123 AspIythrValITPrThrnGlnPheIleIleValMetCylGlyIlePheValaAspThrProLeu 142
       |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

```

Db 1578 ---AAGAGTCATGCGAGATTTTATGACAGCAAGTGAAGAACTCAATAATCAAAAAATA 1634
Qy 143 ThrGluSerIleuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 162
Db 1635 GCAGATATATGACAGGAATAAGCGCTTCGGAATTAGAAAGATTTAGTAATAATTACCA 1694
Qy 163 SerTyrAsnThrAlaLeuAspAspTTParGlyLeuLeuLeuLeuLeuLeuLeuLeuLeu 182
Db 1695 TTATATCTTACTACGCGCTTGAGAAATGGAAGAA----- 1727
Qy 183 ProPheSerSerAlaLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 202
Db 1728 AATCCAAATGCTTCA-----AGAGCTTACGAGATGCGAAATGATTTGAAATTCG 1781
Qy 203 HisAsnAspPheIleArgGluIleProGlyPheGluLeuLeuLeuLeuLeuLeuLeuLeu 222
Db 1782 GATAGTTATTTACGCAATATACATGCACTTTTCAGTGACAAATTTGAGTATACCATTC 1841
Qy 223 LeuProIleTyrAlaGluAlaAlaAsnPheHisLeuAsnLeuLeuLeuLeuLeuLeuLeu 242
Db 1842 CTTACAGTATATACACAGCGACCACTTCATTTTAAAGCAAGCTTCAATTT 1901
Qy 243 LeuAlaAspGluTTPAsnAlaAspIleHisProSerGluIleGluProAsnAlaGlyThr 262
Db 1902 TTTGAGAGAAATGGGAGATG-----TCATCAACCACT 1934
Qy 263 SerAspAspTyrTyrIleLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 282
Db 1935 ATTAATTAATCTATTTATGATCGTCAAAATGAAACTTACGCAAGATATTTCTGATCGTGA 1994
Qy 283 AsnThrTyrArgThrGlyLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 302
Db 1995 AAGTGATGAAACCTGCTTTACCAAAATTTAAAGCACAGCGCTTAAACAAATGCGCTGAC 2054
Qy 303 PheAsnAspTyrArgArgTyrMetThrIleThrValLeuAspThrIleSerGluPheSer 322
Db 2055 TATATCAATTCCTCGTAGAAGAAATGACACTGACGCTTTAGATGTTGTCATATTATCCCA 2114
Qy 323 LeuTyrAspIleLeuArgTyrArgAspSerIleGlyGlyIleGluValIleGlyIleLeu 342
Db 2115 AATTATGACACACGACGCTAC-----CCAATGGAAGAAACGAACA----- 2153
Qy 343 AsnGluLeuThrArgGluIleTyrThrThrGluIleAsnPheAspArgLeuProGluLeu 362
Db 2154 ---CAACTAAACAGGAAGTATATACA-----GATCCACTGGCGCGCTGA 2195
Qy 363 ArgValGlu-----ProAsnLeuAlaThrMetGluTyr 373
Db 2196 AACGTGCTTCATTTGCTTCCTGCTATGACAAAGCACCTTCTTCGAGTGATGAAATCA 2255
Qy 374 AsnLeuThrArgAlaSerPheLeuLeuPheSerPheLeuGluGluPheIlePheTyrThr 393
Db 2256 TCCGTTATTCGACCAACC---CATGATTTTGATTTATATACCGGACCTACAGTGTATACA 2312
Qy 394 GluAsnThrAsnPheGlyAsn-----ArgLeuValGlyIleSerAsnArg 408
Db 2313 CAATCAAGAGCATTTCTCCGCTATATATACAAATTTGGCGTGTCATCAAAATTAAGC 2372
Qy 409 AspAlaProThrTyrSerAsnThrIleThrGluThrLeuTyrGluArgThrGlySer 428
Db 2373 TATCATCGCATTTTATGATATATATTAATAACAGATGATGGAATCAATCAAAATCTA 2432
Qy 429 ProThrThrIleThrIleArgProPheGluSerTyr-----Lys 441
Db 2433 CACAGCAGTACGACTTT---GATTTACGATTTATGATATTTACAGACGTTATCAAAA 2489
Qy 442 ValSerIleValThrAspArgGluSerProPheValSerProIleGluProHisPheIle 461
Db 2490 GATCGGAGTCTCTGATATTTCTTCCTGCTTACGATATA-----TTTTTGCA 2543
Qy 462 IleAsnGluIleGluLeuTyrLeuAsnGlySerSerAsnAsn-----ThrLeuLeu 478
Db 2544 ATGCCAGAAATCGAGTCTTTTCATGAGTAAACCAATTAATAATCAAGAAAGACGTTAAAG 2603

Qy 479 TyrSer-----AlaGlyIleSerLeuSerAsnTyrGluAsnThr 491
Db 2604 TATATCCGCTTCCAAAGATATATATACGGGG----- 2636
Qy 492 ThrPhePheGluPheProArgGlyLeuAspCysAsnLeuValIleAspProGlyCysSer 511
Db 2637 -----ACAAGATTCGGAATTAAGATTTCTTCGAAACCTTCA 2675
Qy 512 -----ProAsnPheAsnTyrSerHisIleLeuSerHisPhe---SerLeuPheThr 528
Db 2676 GATCACCAAAATTAATGATCATATATGCAATTTATGTCATATACAAAGTATTCGCGG 2735
Qy 529 TyrSerTyrValIleGlyLeuGluLeuGluIleLeuAspThrGlyValLeuGlyTyrThr 548
Db 2736 ACGGCTCAACTACCGGATTA-----GTACGCTATTTTCTTGACA 2777
Qy 549 HisSerSerValAspArgTyrAsnAlaIleSerAspGlyIleIleThrMetIleProAla 568
Db 2778 CATCGAGTGGCCGATCTTATTAATGCACTCATTCAGATTAATAATCTCAGATTCGCGTC 2837
Qy 569 IleuGlyAsnAsnLeu-----AspThrAsnSerLeuValIleGlu 582
Db 2838 GTAAAGTCTTGATTTGCTTCCTCTTATACAGAGGCGCAAAATTAATCAGTGTATCG 2897
Qy 583 GlyProGlyHisThrGlyIleAsnLeuValTyrLeuGluSerGluIle----- 598
Db 2898 GGTCTCGATTTTACAGCGGGGGGATATTAAGTAATTAAGTAAGTAAGTATATATCA 2957
Qy 599 -----ArgLeuGluIleThrCysGluThrProAsnSerThrGluSerTyrPheIleArg 616
Db 2958 CATATGCGTGTAAATTTCA-----GACTTAACAAAGATTAATGATGTAGAG 3005
Qy 617 LeuArgTyrAlaThrAsnGlyValGlyAsnThrLeuProAsnIleSerLeuThrIlePro 636
Db 3006 ATTCGATTCGCTTC-----GCTATATATCTGAATTTTATATATATCTTCTTAAGAA 3059
Qy 637 GlyValIleGlyIleProProGluArgLeuAsnAsnThrPheSerGlyThrAsnTyrAsn 656
Db 3060 AACGTTAAATCT---CACGCTCAAAAACSTAGTAATAGAGGAAAGCTTAAACATATAT 3116
Qy 657 AsnLeuGluTyrGlyAspPheGlyTyrPheGluPheProSerThr-----ValThr 673
Db 3117 AAATTAATTAATGACACTTTCGCTTATTAATTAATTAATTAATTAATTAATTAAT 3176
Qy 674 LeuProLeuAsnArgAsnIleProPheIlePheAsnArgAlaAspValSerAsnSerIle 693
Db 3177 CTAGGGGCT-----ATATTTGAAGCGGAAGACTTTCTTGAAATTGAA 3218
Qy 694 LeuIleIleAspValIleGluPheIleProIleThrSerSerMetHisGluAsnArgGlu 713
Db 3219 GCTTATATAGACCAAGATGATTTATCCCAATGATGACATAT-----GAACCG 3269
Qy 714 LysGluLeuLeuGluThrIleGluThrIleLeuAsnThrPhePheThrAsn 730
Db 3270 GAACAGATTTTGAAGACAGCAAGAAAGAGTGAATGCTTGTATACGAAT 3320
RESULT 14
US-10-414-637-28
; Sequence 28, Application US/10414637
; Publication No. US20030177528A1
; GENERAL INFORMATION:
; APPLICANT: Andre R. Adad
; APPLICANT: Nicholas B. Duck
; APPLICANT: Xiang Feng
; APPLICANT: Ronald D. Flannagan
; APPLICANT: Theodore W. Kahn
; APPLICANT: Lynn R. Sims
; TITLE OF INVENTION: Genes Encoding No. US20030177528A1el Proteins With
; TITLE OF INVENTION: Pesticidal Activity Against Coleopterans
; FILE REFERENCE: 35716/237005
; CURRENT APPLICATION NUMBER: US/10/414, 637
; CURRENT FILING DATE: 2003-04-16

PRIOR APPLICATION NUMBER: US/10/032,717
PRIOR FILING DATE: 2001-10-23
PRIOR APPLICATION NUMBER: 60/242,838
PRIOR FILING DATE: 2000-10-24
NUMBER OF SEQ ID NOS: 48
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 28
LENGTH: 6613
TYPE: DNA
ORGANISM: *Bacillus thuringiensis*
FEATURE:
NAME/KEY: misc_feature
LOCATION: (0)...(0)
OTHER INFORMATION: Genomic Cry121a-2
US-10-414-637-28

Alignment Scores:
Pred. No.: 1,47e-63 Length: 6613
Score: 703.00 Matches: 219
Percent Similarity: 44.5% Conservative: 127
Best Local Similarity: 28.2% Mismatches: 289
Query Match: 18.2% Indels: 142
Gaps: 32
DB:

US-10-783-417-2 (1-735) x US-10-414-637-28 (1-6613)

QY 5 AaapAspAsnAgLuTrpGluIleIleAspSerHisThrSerProTyrrPheProAsnArg 24
DB 1263 AATAAATCAAAATGAAATATGAAATATGATGACGACACTTCT-----ACTTCT 1310
QY 25 AaenSerAsnAspSer---ArgTyrrProTyrrThraAsnProAsnGlnProLeuGlnAan 43
DB 1311 GATATCAATGATTTCTAAGATACCTCTTTGGCAATGAGCCAAATGCGCTACAAAT 1370
QY 44 ThrAenTyrrLeuGluTrpLeuAsnMetCysGlnGlyAsn---ThrlGlnTyrrGlyAspAan 62
DB 1371 ATGGAATTAATAAGATATTAATAAATGCTGCGGAAATGCTGATGATATCCTGGTCA 1430
QY 63 PheGluThrPheAlaSerAlaPheThrIleAlaIleValSerAlaGlyThrIleValSer 82
DB 1431 CCTGAAGTACTGTTAGCGGACAAAGAT-----GCAAGCTAAGGCCCAATTGATATAGA 1484
QY 83 GlyThrLeuLeuAlaGlyIleGlyIleLeuThrSerIleSerGlyProIleGlyIleIle 102
DB 1485 GGTAAATTAATCTATCAGGTTTAAAGG-----GTCCATTGTT 1520
QY 103 GlyAlaIleIleIleSerPheGlyThrLeuIleThrValPheTyrrProAlaGlyGln 122
DB 1521 GGGCCGATAGTAGTCTTATATCTCAACTATGATATCTGCGCTTCAAGGCAA--- 1577
QY 123 AspIleThrValIleThrGlnPheIleIleValMetGlyIleIlePheValAspThrProLeu 142
DB 1578 ---AAAGTCATGCGAGATTTTATGAAACAAGTAAAGACCTTAATCAAAAATA 1634
QY 143 ThrGluSerIleValGlnLeuIleValGlnThrLeuGlnGlyPheArgGlnIleLeuGln 162
DB 1635 GCGAATATGACGAGAAATTAAGCGCTTTCGGAATTAAGAGATTAAGTAATTAATCCAA 1694
QY 163 SerTyrrAsnThrAlaLeuAspAspTyrrArgIleValLeuIleValArgLeuGlnIleProGlyLeu 182
DB 1695 TTATATCTAATCGCGCTTGAAGATGAAAGAA----- 1727
QY 183 ProProSerSerAlaLeuGlnGlnAlaIleLeuThrLeuIleValIleArgPheGluAsnVal 202
DB 1728 AATCCAAATGCTTCA-----AGAGCCTTAAGAGATGCGAATGATTTGAAATCCG 1781
QY 203 HisAsnAspPheIleArgIleIleProGlyPheGlnLeuGlnThrTyrrIleThrLeuLeu 222
DB 1782 GATAGTTTATTTACGAATATGATGCAATCTTTTCGATGACAAATTTGAGTACCAATTC 1841
QY 223 LeuProIleTyrrAlaGlnAlaIleAsnPheHisIleLeuAsnLeuLeuGlnGlnAlaGln 242
DB 1842 CTTACAGTATATACAGGCAAGCCAACTTCATTTACTGTTATTAAGGACGCTTCAAT 1901

QY 243 LeuAlaAspGluTrpAsnAlaAspIleHisProSerGlnIleGluProAsnAlaGlyThr 262
DB 1902 TTTCGAGAAAGATGGGATGG-----TCTAACACACT 1934
QY 263 SerAspAspTyrrTyrrIleValLeuIleValGlnIleProIleTyrrSerAsnTyrrCysAla 282
DB 1935 ATTAATTAATTAATTAATGATGATGCAAAATGAATGAACTTACTGACAAATATCTGATCCTGTGA 1994
QY 283 AsnThrTyrrArgThrGlyLeuIleValAsnLeuAlaGlyAspGluProAsnMetIleTyrrSerIle 302
DB 1995 AAGTGATAGAAATCGGTTTACCAAAATTAAGCAAGCGCTTAACAAATGGCTGAC 2054
QY 303 PheAsnAspTyrrArgArgTyrrMetThrIleThrValIleAspThrIleSerGlnPheSer 322
DB 2055 TATTAACCAATTCGATGAGAAATGACACCTGACGCTTTTGAATGTTGTTCGATTAATCCCA 2114
QY 323 LeuTyrrAspIleValArgTyrrArgAspSerIleGlyIleGluValIleGlyIleVal 342
DB 2115 AATTATGACACACGACGCTAC-----CCAATGAAACGAAAGCA----- 2153
QY 343 AsnGluLeuThrArgGluIleTyrrThrGlnIleAsnPheAspArgLeuProGlnLeu 362
DB 2154 ---CAACTAACAGGAGAGATATATACA-----GATCACTGGCGCGGCTA 2195
QY 363 ArgValGln-----ProAsnLeuAlaThrMetGluTyrr 373
DB 2196 AACGTTCTTCAATGTTCTTCGCTGATGACAAAGCACTTCTTCGAGATGATGAATCA 2255
QY 374 AsnLeuThrArgAlaSerPheIleValLeuPheSerPheLeuGlnGlnPheIlePheTyrr 393
DB 2256 TCGGTTATTCGACCAACC---CATGATTTGATTAATAACGGGACTCACAGTATATAC 2312
QY 394 GluAsnThrAsnPheGlyAsn-----ArgLeuValGlyIleSerAsnArg 408
DB 2313 CAATCAAGAAAGATTTCTTCGCTGATTAATAAGCAATTTGGCTGATCAAAATAGC 2372
QY 409 AspAlaProThrTyrrSerAsnThrIleThrGluThrLeuTyrrGlyValuArgThrGlySer 428
DB 2373 TATCATCGGATTTTATGATATATTAATAACAGATATAGAACTAATCAAAATCTA 2432
QY 429 ProThrThrIleThrIleArgProPheGluSerTyrr-----Lys 441
DB 2433 CACAGACCTAGTACCTT---GATTTTACGATTAATGATTTACAGACGTTATGAAA 2489
QY 442 ValSerIleValIleThrAspArgIleSerProValSerProIleGlnProHisPheIle 461
DB 2490 GATCGGCTGCTCTGATGATGTTTCTTCTGTTATACGATATATA-----TTTTTGA 2543
QY 462 IleAsnGlnIleGluLeuTyrrIleAsnGlnIleSerSerAsnAsn-----ThrLeuIle 478
DB 2544 ATGCCAAGATGCAAGTTTTCATGTGAACCAATTGAATAAATCCAGAAAGACGTTAAG 2603
QY 479 TyrSer-----AlaGlyGlySerLeuSerAsnTyrrGlnAsnThr 491
DB 2604 TATATCCGTTTCCAAAGATATTAATAGCGGG----- 2636
QY 492 ThrPhePheGlnPheProArgIleValAspCysAsnLeuValIleAspProGlyCysSer 511
DB 2637 -----ACAAGAGATTCGGAATTAAGATTAATCACTCCAGAAACTTCA 2675
QY 512 -----ProAsnPheAsnAsnTyrrSerHisIleLeuSerHisPhe---SerLeuPheThr 528
DB 2676 GATCAACAAATTAATGAGTATAGCCATTAATGATGATATGATATCAAGATATTCGCGG 2735
QY 529 TyrSerTyrrValIleGlyLeuGlnLeuGlnIleLeuAspThrGlyValLeuGlyTyrrThr 548
DB 2736 ACGGTTCACTACCGGATTA-----GTACCTGATATTTCTTGACA 2777
QY 549 HisSerSerValAspArgTyrrAsnAlaIleSerAspIleIleThrMetIleProAla 568
DB 2778 CATCGAGTGGCGGATCTTATTAATGACAGTTCATTCAGATTAATAATTAATCAATTCGCGT 2837

QY 569 ILELYGILYAsnAsnLeu-----AspThrAsnSerIyValIleGlu 582
 Db 2838 GTAAGGTTCTGATTTGGCTCCCTATACAGAGGCGCAAAATATACGTTGATCG 2897
 QY 583 GTPGPGIYHISThGILYAsnLeuValIYrLeuGlnSerGlnGly----- 598
 Db 2898 GGTCTCGATTTACAGGAGGAGTAATAAGATATAGAAATGAGATTAATATATCA 2957
 QY 599 -----ArgLeuGluIleThrCysGluThrProAsnSerThrGlnSerIyPheIleArg 616
 Db 2958 CATACCGGTTAAATTTCA-----GACATTAACAAGAAATATAGTATAGG 3005
 QY 617 LeuArgTYrAlaThrAsnGlyAlaGlyAsnThrLeuProAsnIleSerLeuThrIlePro 636
 Db 3006 ATTCGGTATGCTTCC-----GCTATATATACGATTTTATATATATCTCTGAAGA 3059
 QY 637 GLYValIleGlyIYrLeuProGlnArgLeuAsnAsnThrPheSerGlyThrAsnTYrAsn 656
 Db 3060 AAGGTAAATCT--CACGCTCAAAAACCTATGATGAGTGAAGCTTTAACATATAT 3116
 QY 657 AsnLeuGlnTYrGlyAspPheGlyTYrPheGlnPheProSerThr-----ValThr 673
 Db 3117 AATTTAATTAATTCGCACTTGGCCCTATTAATTTACGACAAACGACCTTTCATTA 3176
 QY 674 LeuProLeuAsnArgAsnIleProPheIlePheAsnArgAlaAspValSerAsnSerIle 693
 Db 3177 CTAGGGGCT-----ATATTTGAACGCGAAGACTTCTTGGAAATTGAA 3218
 QY 694 LeuIleIleAspIyIleGluPheIleProIleIleThrSerSerMetHisGlnAsnArgGlu 713
 Db 3219 GCTTATATAGACCGAATTCGAATTTATCCAGTATGATGAGACATAT-----GAAAGC 3269
 QY 714 LysGlnLysLeuGluIleGlnThrIleGlnThrIleAsnThrPhePheThrAsn 730
 Db 3270 GAACAAGATTATGAAACGACGAAAGAAAGACGTAATGCTTGTTCGAT 3320

RESULT 15
 US-10-606-320-18
 ; Sequence 18, Application US/10606320
 ; Publication No. US20040091505A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Andre R. Abad
 ; APPLICANT: Ronald D. Flannagan
 ; APPLICANT: Rafael Herrmann
 ; APPLICANT: Theodore W. Kahn
 ; APPLICANT: Albert L. Lu
 ; APPLICANT: Billy Fred McCutchen
 ; APPLICANT: James K. Presnail
 ; APPLICANT: James F.H. Wong
 ; APPLICANT: Cao-Guo Yu
 ; TITLE OF INVENTION: Genes Encoding Proteins With Pesticidal
 ; FILE OF INVENTION: Activity
 ; FILE REFERENCE: 35718/263948
 ; CURRENT APPLICATION NUMBER: US/10/606,320
 ; PRIOR FILING DATE: 2003-06-25
 ; PRIOR APPLICATION NUMBER: 60/391,786
 ; PRIOR FILING DATE: 2002-06-26
 ; PRIOR APPLICATION NUMBER: 60/460,787
 ; PRIOR FILING DATE: 2003-04-04
 ; NUMBER OF SEQ ID NOS: 134
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 18
 ; LENGTH: 6613
 ; TYPE: DNA
 ; ORGANISM: Bacillus thuringiensis
 ; FEATURE:
 ; NAME/KEY: misc_feature
 ; LOCATION: (0)...(0)
 ; OTHER INFORMATION: Genomic Cyt1218-2
 ; US-10-606-320-18
 Alignment Scores:
 Pred. No.: 1.47e-63 Length: 6613

Score: 703.00 Matches: 219
 Percent Similarity: 44.5% Conservative: 127
 Best Local Similarity: 28.2% Mismatches: 289
 Query Match: 18.2% Indels: 142
 DB: 7 Gaps: 32
 US-10-783-417-2 (1-735) x US-10-606-320-18 (1-6613)
 QY 5 AsnAspAsnAsnGluTYrGluIleIleAspSerHisThrSerProTYrPheProAsnArg 24
 Db 1263 AATATCAAAATGATATGAAATATTAATTAATGACGACACCTTCT-----ACTTCT 1310
 QY 25 AsnSerAsnAspSer--ArgTYrProTYrThrAsnAsnProAsnGlnProLeuGlnAsn 43
 Db 1311 GTATCAATATGATTCATACGATATCCCTTTGGGAATGAGCCCAAAATGCGCTACAAAAT 1370
 QY 44 ThrAsnTYrLysGluIleThrLeuAsnMetCysGlnGlyAsn--ThrGlnTYrGlyAspAsn 62
 Db 1371 ATGATATTAATAAGATTTATTAATAATGCTCGGGAATAATGCTATGATATACCTGGTTCA 1430
 QY 63 PheGluThrPheAlaSerAlaAspThrIleAlaValaValaSerAlaGlyThrIleValaSer 82
 Db 1431 CCTGAAGTACTGTTATGCGGACAAAGAT-----GACGCTAAGCGCGCAATATATATAGA 1484
 QY 83 GlyThrLeuLeuAlaGlyIYrLeuGlyLeuThrSerIleSerGlyProIleGlyIleIle 102
 Db 1485 GGTAAATTAATCTATCAGTTAGGG-----GTCCCATTTGTT 1520
 QY 103 GYValIleIleIleSerPheGlyTYrLeuIleThrValaPheThrProAlaGlyGluGln 122
 Db 1521 GGGCGGATGATGAGTCTTTATATCTCAACTATTTGATATTTCTGTCGCTTACAGGCA-- 1577
 QY 123 AspIyThrValaTYrThrGlnPheIleLysMetGlyGluIlePheValaAspThrProLeu 142
 Db 1578 ---AAGATCAATGGAAGATTTTATGCAACAAGATGAAAGAACTGATTAATCAAAAAATA 1634
 QY 143 ThrGluSerIleLysGlnLeuLysLeuGlnThrLeuGlnGlyPheArgGlnIleLeuGln 162
 Db 1635 GCAGATATGCAAGGAATTAAGCGCTTCCGAAATTAGAAGATAGTATATATATACCA 1694
 QY 163 SerTYrAsnThrAlaLeuAspAspTYrParGlyLysLeuLysArgLeuGlnAlaProGlyLeu 182
 Db 1695 TTATATCTAATCGCCCTTGAAAGAAATGAAAGAA----- 1727
 QY 183 ProProSerSerAlaLeuGlnGlnAlaIleLeuThrIleLysIleArgPheGluAsnVal 202
 Db 1728 AATCCAATGCTTCA-----AGAGCTTACGAGATGCGCAAAATCGATTAATATCCG 1781
 QY 203 HisAsnAspPheIleArgGluIleProGlyPheGlnLeuGluIleThrTYrLysThrLeuLeu 222
 Db 1782 GATAGTTATTTACCAATATACATGCAATCTTTTCCAGTACGACAAATTTTGAAGTACCATTC 1841
 QY 223 LeuProIleTYrAlaGlnAlaAlaAsnPheHisLeuAsnLeuLeuGlnGlnGlyAlaGlu 242
 Db 1842 CTTACAGTATATACACAGGACCACTTCACTTATCTGTATTAAGAGCCTTCAATT 1901
 QY 243 LeuAlaAspGluIleTrpAsnAlaAspIleHisProSerGlnIleGluProAsnAlaGlyTYr 262
 Db 1902 TTTGAGAAAGATGGGATGC-----TCTAACACCAT 1934
 QY 263 SerAspAspTYrTYrLysLeuLeuLysGluAsnIleProLysTYrSerAsnTYrCysAla 282
 Db 1935 ATTAATATCTATTAATGATGCTCAATATGAATTAATCTGACGATATTTCTGATCATCTGTGA 1994
 QY 283 AsnThrTYrArgThrGlyLeuLysAsnLeuAlaGlyAspGluProAsnMetLysTrpSerIle 302
 Db 1995 AAGTGATATGAACCTGTTTACCAAAATTAAGACACAGAGGCTAAACAAATGGGTGCAC 2054
 QY 303 PheAsnAspTYrArgTYrMetThrIleThrValaIleAspThrIleSerGlnPheSer 322
 Db 2055 TATTAACCAATTCGATGAGAAATGACACTGACGGTTTATGATGTTGTTGATTAATTCCTCA 2114
 QY 323 LeuTYrAspIleLysArgTYrArgAspSerIleGlyGlyIleGluValaLysGlyIleLys 342

THIS PAGE BLANK (US#T0)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - nucleic search, using frame_plus.p2n model

Run on: January 20, 2006, 17:08:39 ; Search time 334 Seconds
(without alignments)
3911.700 Million cell updates/sec

Title: US-10-783-417-2

Perfect score: 3869
Sequence: 1 MNQNNNDNNVEYILDSHTSPY.....KLETIQTQKINFTNHTKTL 735

Scoring table:

BLOSUM62	
Xgapop 10.0 , Xgapext 0.5	
Ygapop 10.0 , Ygapext 0.5	
Fgapop 6.0 , Fgapext 7.0	
Delop 6.0 , Delext 7.0	

Searched: 1303057 seqs, 888780828 residues

Total number of hits satisfying chosen parameters: 2606114

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODE=frame+g2n.model -DEV=xlh
-O=/cg2n_1/USPTO_spool/US10783417/runat_20012006_095011_23893/app_query.fasta_1.903
-DB=Issued Patents NA -QFMT=fastp -SUFFIX=g2n.rml -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=biosum62 -TRANS=humana0.cdi
-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=pct -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US10783417 -CGN_1_1_200 @runat_20012006_095011_23893 -NCPU=6 -ICPU=3
-NO MMAP -NMG SCORES=0 -WAIT -DSPBLOCK=100 -DEV TIMEOUT=120 -WARN TIMEOUT=30
-THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7 -YGAPOP=10
-YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Issued Patents_NA:*
1: /cg2n_6/ptodata/1/ina/1_COMB.seq:*
2: /cg2n_6/ptodata/1/ina/5_COMB.seq:*
3: /cg2n_6/ptodata/1/ina/6A_COMB.seq:*
4: /cg2n_6/ptodata/1/ina/6B_COMB.seq:*
5: /cg2n_6/ptodata/1/ina/H_COMB.seq:*
6: /cg2n_6/ptodata/1/ina/PCTUS_COMB.seq:*
7: /cg2n_6/ptodata/1/ina/PP_COMB.seq:*
8: /cg2n_6/ptodata/1/ina/RE_COMB.seq:*
9: /cg2n_6/ptodata/1/ina/Backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1060.5	27.4	3543	3	US-09-224-024-27
2	1060.5	27.4	3543	3	PCT-US94-07902-27
3	899.5	23.2	2061	3	US-09-224-024-30
4	899.5	23.2	2061	6	PCT-US94-07902-30
5	775.5	20.0	3507	2	US-08-315-468-3
6	676	17.5	1953	2	US-08-315-468-5
7	676	17.5	1953	3	US-07-941-650A-3
8	663.5	17.1	2430	6	PCT-US92-00040-1
9	663	17.1	1959	3	US-08-996-441B-37

10	663	17.1	1959	3	US-08-993-722A-37	Sequence 37, Appl
11	663	17.1	1959	3	US-08-993-170A-37	Sequence 37, Appl
12	663	17.1	1959	3	US-08-993-775B-37	Sequence 37, Appl
13	663	17.1	1959	3	US-09-427-770-37	Sequence 37, Appl
14	663	17.1	1959	3	US-09-427-769-37	Sequence 37, Appl
15	660.5	17.1	1959	3	US-07-973-320-3	Sequence 3, Appl1
16	660	17.1	1959	3	US-09-377-466B-3	Sequence 3, Appl1
17	660	17.1	1959	3	US-10-232-665-3	Sequence 3, Appl1
18	658	17.0	1959	3	US-08-996-441B-37	Sequence 27, Appl
19	658	17.0	1959	3	US-08-996-441B-37	Sequence 27, Appl
20	658	17.0	1959	3	US-08-993-722A-27	Sequence 27, Appl
21	658	17.0	1959	3	US-08-993-722A-35	Sequence 35, Appl
22	658	17.0	1959	3	US-08-993-170A-27	Sequence 27, Appl
23	658	17.0	1959	3	US-08-993-170A-35	Sequence 35, Appl
24	658	17.0	1959	3	US-08-993-775B-35	Sequence 35, Appl
25	658	17.0	1959	3	US-08-993-775B-35	Sequence 35, Appl
26	658	17.0	1959	3	US-09-427-770-27	Sequence 27, Appl
27	658	17.0	1959	3	US-09-427-770-35	Sequence 35, Appl
28	658	17.0	1959	3	US-09-427-769-35	Sequence 35, Appl
29	658	17.0	1959	3	US-09-427-769-35	Sequence 35, Appl
30	656	17.0	1962	3	US-09-377-466B-5	Sequence 5, Appl1
31	656	17.0	1962	3	US-10-232-665-5	Sequence 5, Appl1
32	656	17.0	2050	3	US-08-996-441B-101	Sequence 101, App
33	656	17.0	2050	3	US-08-993-722A-101	Sequence 101, App
34	656	17.0	2050	3	US-08-993-170A-101	Sequence 101, App
35	656	17.0	2050	3	US-08-993-775B-101	Sequence 101, App
36	656	17.0	2050	3	US-09-427-770-101	Sequence 101, App
37	656	17.0	2050	3	US-09-427-769-101	Sequence 101, App
38	655	16.9	1959	3	US-08-996-441B-17	Sequence 17, Appl
39	655	16.9	1959	3	US-08-996-441B-59	Sequence 59, Appl
40	655	16.9	1959	3	US-08-993-722A-17	Sequence 17, Appl
41	655	16.9	1959	3	US-08-993-722A-59	Sequence 59, Appl
42	655	16.9	1959	3	US-08-993-170A-17	Sequence 17, Appl
43	655	16.9	1959	3	US-08-993-170A-59	Sequence 59, Appl
44	655	16.9	1959	3	US-08-993-775B-17	Sequence 17, Appl
45	655	16.9	1959	3	US-08-993-775B-59	Sequence 59, Appl

ALIGNMENTS

RESULT 1
US-09-224-024-27
Sequence 27, Application US/09224024
Patent No. 6056953
GENERAL INFORMATION:
APPLICANT: Leslie Hickie
TITLE OF INVENTION: Materials and Methods for the Control of
TITLE OF INVENTION: Calliphoridae Pests
NUMBER OF SEQUENCES: 31
CORRESPONDENCE ADDRESS:
ADDRESSEE: David R. Saliwanichik
STREET: 2421 N.W. 41st Street, Suite A-1
CITY: Gainesville
STATE: FL
COUNTRY: USA
ZIP: 32606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/224,024
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/856,226
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Saliwanichik, David R.
REGISTRATION NUMBER: 31,794
REFERENCE/DOCKET NUMBER: MA79

[illegible][illegible]


```

Db      784 AATTATGTGTGAACAACCTTATATAAAAGATTAAATTTAATTAAACGACGCTGATAGT 843
Oy      297 -----AsmMetIytrpSerIlePheAsnAapTyArgArgTyserThrIle 312
Db      844 AATCTTGATGAAATATATACTGGAACACATACATACATGATCGAACAAAATGACTCT 903
Oy      313 ThrValIleuAapThrIleSerGlnPheSerLeuTyArgPleuArgTyArgAapSer 332
Db      904 GCTGATATAGATCTTGACTCTTCTCTAATATGATGATGATGATGATGATGATGAT 954
Oy      333 IleGlyGlyIleGluValIleGlyIleuAsnGlnLeuThrArgGluIleTyThrThr 352
Db      955 -----CCAATAGTGTCCAAATCTGAACCTTACCGAATAATTATCAGGTA 999
Oy      353 GluIleAsnPheAapArgLeuProGlnLeuArgValGlnProAsnLeuAlaThrMetGlu 372
Db      1000 -----CTTAACCTTGAAAGAAAGCCCTTAATAATTAT-----GACTTCGAATATCAAGG 1050
Oy      373 TyrAsnLeuThrArgAlaSerPheLeuPheSerPheLeuGlnPheIlePheTy 392
Db      1051 GATTCACCTTACACGTAGACCG---CATTTATTTACTTGCTGATTCCTTGAAATTTTAT 1107
Oy      393 -----ThrGluAsnThrAsnPheGlyAsnArg----- 401
Db      1108 GAAAAAGCGCAACTACTCTTAATATTTTTCACACGCACTTAATATATGTTTCATTAC 1167
Oy      402 ---LeuValIleGlyIleSerAsnArgAapAlaProThrTySerAsnThrIleThrGluThr 420
Db      1168 ACACCTTGATATATATATCCAAAATCTAGTGTGTTTGGAAATCAACAATGATACATATAA 1227
Oy      421 LeuTyArgIleuArgTyArgTyserProThrThrIleArgProPheGluSerTy 440
Db      1228 TTA-----AATCTCTGTTGGTGGCAACAATATTATATT----- 1263
Oy      441 LysValSerIleValThrAapArgIleSerProProValSerProIleGlnProHisPhe 460
Db      1264 -----TTT 1266
Oy      461 IleIleAsnGlnIleGluLeu-----TyrLeuAsnGlySerSerAsnAsnThr--- 476
Db      1267 TTATTAATATGTCATTAAGCTTATAGATATAATATCTTAATATGATTATATATTAATA 1326
Oy      477 -----LeuTySerIleGlyIle 483
Db      1327 ATGATTTTATATACTAATAGTACTAGCTTTTGGAAAGAACTTACACAGCAAGACT 1386
Oy      484 SerLeuSerAenTyGln---AsnThrThrPhePheGlnPheProArgIleLysAapCys 502
Db      1387 GGGCAAAATTAACCTTATGATGATTAATAAATATTTTCGGGTTTCAATCTTAAACCAAGA 1446
Oy      503 AsnLeuValIleAapProGlyCysSerProAsnPheAsnAenTySerHisIleLeuPhe 522
Db      1447 GAGAAATCAAGCAATGCTTACCTTTTCCACATATGATACATATATATATATATATCA 1506
Oy      523 HisPheSerLeuPheThrTySerTyValIleGlyLeuGlnLeuGlnIleLeuAapThr 542
Db      1507 TTTATTAAGAAGCTTAGTATCCCTGCAACATATAAAGCAAGTGAT----- 1554
Oy      543 GlyValIleuGlyTyrThrHisSerSerValAapArgTyArgAlaIleSerAapLysIle 562
Db      1555 ---ACGTTTGCTTGACACACTCTAGTGTGATCTTAATAAATACATTTATATACACATTYA 1611
Oy      563 IleThrMetIleProAlaIleLeuGlyAsnAsnLeuAapThrAsnSerLysValIleGlu 582
Db      1612 ACTACCCAAATTCAGCTGATTAAGCGAATTCAGTGGACGCTTCTTAAGGTATTTCA 1671
Oy      583 GlyProGlyHisThrGlyGlyAsnLeuValTyLeuGlnSerGlnGlyArgLeuGluIle 602
Db      1672 GGAAGCTGGTGCATACAGAGGGGATTTAAT-----GATTCAAGATCATTTCAAAATT 1725
Oy      603 ThrCysGluThrProAsnSerThrGlnSerTyPheIleArgLeuArgTyArgAlaThrAsn 622

```

```

Db      1726 ACATGTCACACCTCAAAATTTTCAACATCGTATTTTATAGAAATTCGTTAGCTTCAAAAT 1785
Oy      623 GlyAlaGlyAsnThrLeuProAsnIleSerLeuThrIleProGlyValIleGlyIlePro 642
Db      1786 GGAAGCGCAAAATACACAGCGCTGTTATTAATCTTATGATCCAGGGGTGACGAAGCTG--- 1842
Oy      643 ProGlnArgLeuAsnAenThrPheSerGlyThrAsnTyArgAsnAsnLeuGlnTyGlyAap 662
Db      1843 GGTATGACCACTCAACCCCACTTTTCTGTCAGATTTATAGAAATTTAAATATATGAAGAT 1902
Oy      663 PheGlyTyPheGlnPheProSerThrValThrLeuProLeuAsnArgAsnIleProPhe 682
Db      1903 TTTCAAGTACTTAAGATTTTCTTAACGAGGTGAATTTGCTCCAAATCAAAACATATCTCTT 1962
Oy      683 IlePheAsnArgAlaAapVal---SerAsnSerIleLeuIleIleAapLysIleGluPhe 701
Db      1963 GTGTTTAATCGTTTCGATGATATATACAAACACACAGTACTTATGATTAATTAATTT 2022
Oy      702 IleProIleThrSerSerMetHisGlnAsnArgIleuGlyGlyLeuGluThrIleGln 721
Db      2023 CTGCAATTTACTCGTCTATTAAGAGAGATGAGAGAAACAAAATTTGAAAACAGTACAA 2082
Oy      722 ThrLysIleAsnThrPhePheThrAsnHisThrLys 733
Db      2083 CAATTAATTAATACATTTTATGCAAAATCTATATAAA 2118

RESULT 3
US-09-224-024-30
; Sequence 30, Application US/09224024
; Patent No. 6056953
; GENERAL INFORMATION:
; APPLICANT: Leslie Hickie
; APPLICANT: Jewel Payne
; TITLE OF INVENTION: Materials and Methods for the Control of
; TITLE OF INVENTION: Calliphoridae Pests
; NUMBER OF SEQUENCES: 31
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: David R. Saliwanchik
; STREET: 2421 N.W. 41st Street, Suite A-1
; CITY: Gainesville
; STATE: FL
; COUNTRY: USA
; ZIP: 32606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/224, 024
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/856,226
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Saliwanchik, David R.
; REGISTRATION NUMBER: 31,794
; TELEPHONE: 904-375-8100
; TELEFAX: 904-375-8100
; REFERENCE/DOCKET NUMBER: MA79
; TELECOMMUNICATION INFORMATION:
; INFORMATION FOR SEQ ID NO: 30:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2061 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; US-09-224-024-30

Alignment Scores: 1.61e-85 Length: 2061
Pred. No.:

```

Score: 899.50 Matches: 247
Percent Similarity: 48.3% Conservative: 126
Best Local Similarity: 32.0% Mismatches: 270
Query Match: 23.2% Indels: 129
DB: 3 Gaps: 30

US-10-783-417-2 (1-735) x US-09-224-024-30 (1-2061)

QY 1 MetAenGlnAsnAsnAsnAsnAsnAsnGluTyrGluIleIleAspSerHisThrSerProTyr 20
DB 1 ATGATCCCATATCAAAATAGATGAATGAAATATGAAATATTCAGATGCTCATCAATGGTTT 60
QY 21 PheProAsnArgAsnSerAsnAsnSerAsnSerArgTyrProTyrThrAsnAsnProAsnGlnPro 40
DB 61 -----AGCAAGCTAATACTATCTAGATATCCATTAGCAAAATAGCAAAATCAACA 114
QY 41 LeuGlnAsnThrAsnTyrLysGluTyrPLeuAsnMetCysGlnGlyAsnThrGlnTyrGly 60
DB 115 CTGAAAAACACGAAATACAAAGATTGGCTCAATGTGTCAGATATCAAGATATCAACATATGCG 174
QY 61 AspAsnPheGluThrPheAlaSerAlaAspThrIleAlaAlaValSerAlaGlyThrIle 80
DB 175 AATTAATGCGGGAAATTTTGTAGTCTGAAACTATTTGTGAGTTAGTCAGGTATATT 234
QY 81 ValSerGlyThrLeuLeuAlaGlyIleGlyGlyLeuThrSerIleSerGlyProIleGly 100
DB 235 GTAGTAGGAAGTATGTTAGGA-----GCTTTTCCTGCCCC 270
QY 101 IleIleGlyAlaIleIleIleSerPheGlyThrLeuIleThrValPheTyrProAlaGly 120
DB 271 GCTTAGTCGACAGGTATATATCTTTTGGGACTTTTGGCCATCTTTGG---CAGGA 327
QY 121 GluGlnAspLysThrValTyrThrGlnPheIleLysMetGlyGluIlePheValAspThr 140
DB 328 TCTGACCTGCAAAATGTTGGCAGATTGTTAAACATCGA-----GGAGG 375
QY 141 ProLeuThrGlu-----SerIleLysGlnLeuLysLeuGlnThrLeuGlnGlyPhe 157
DB 376 CTTATTCAGAAATATAGATATAAAACATATATATATGTTCTTAATCTGTAACACTATA 435
QY 158 ArgGlnIleLeuGlnSerTyrAsnThrAlaLeuAspAspTyrArgLysLeuLysValGlu 177
DB 436 AAAAATCAACCTGATTAATATCAAGAAATTTTTCGATTAATGGAGCCAGACAGCT----- 489
QY 178 GlnAlaProGlyLeuProProSerSerAlaLeuGlnGlnAlaAlaLeuThrLeuLysIle 197
DB 490 -----ACACACGCTAATGCTTAAGCAGTACATGATCTCTTACT 528
QY 198 ArgPheGluAsnValHisAsn-----AspPheIleArgGluIleProGlyPheGln 214
DB 529 ACCTTAGAACCTTAATATAGATTAAGATTAGATATGTTAAATAATATGCTAGCTATGCA 588
QY 215 LeuGluThrTyrLysThrLeuLeuLeuProIleTyrAlaGlnAlaAlaAsnPheHisLeu 234
DB 589 ATACCAACA-----CTCCCTGCATATGCACAAATAGCTACTTGGCCTTG 633
QY 235 AsnLeuLeuGlnGlnGlyAlaGluLeuAlaAspGluTyrAsnAlaAspIleHisProSer 254
DB 634 AATTTATTTAAACATGCTGCTACTACTTACCAATATATG-----CTGCAAAATCA 684
QY 255 GlnIleGluProAsnAlaGlyThrSerAspAspTyrTyrLys-----LeuLeuLysGluAsn 273
DB 685 GGTATTAATCAAGTACTTCAATTCATCTTAATTAATACATGAGGCTAATTTAAACGTAA 744
QY 274 IleProLysTyrSerAsnTyrCysAlaAsnThrTyrArgThrGlyLeuLysAsnLeuArg 293
DB 745 ATACAGAAATATATCTGATTTGTATACAAAGCTACAAAGCAGCACTAATGATGATTAG 804
QY 294 AspGluProAsnMetLysTyrSerIlePheAsnAspTyrArgArgTyrMetThrIleThr 313
DB 805 ACTAATATACATGCAACATGGAATATGTTAATATCTTACCGTTTACAAATGACTCTAACT 864
QY 314 ValLeuAspThrIleSerGlnPheSerLeuTyrAspIleLysArgTyrArgAspSerIle 333

DB 865 GTGTAGATCTTATTCGATATTTTCCAAATATATGACCCAGAAATAT----- 912
QY 334 GlyGlyIleGluValLysGlyIleLysAsnGluLeuThrArgGluIleTyrThrGlu 353
DB 913 -----CCAAATGAGTAAATCTGAACTTACCAAGAGATTTAT---ACGAAT 957
QY 354 IleAsnPheAspArgLeuProGlnLeuArgValGlnProAsnLeuAlaThrMetGluTyr 373
DB 958 GTTAATTCAGTACATTT-----AGAACATAAACAGAACTAGAAAT 999
QY 374 AsnLeuThrArgAlaSerPheLysLeuPheSerPheLeuGlnGlnPheIlePheTyrThr 393
DB 1000 GGATTAATCTAGA---AATCTCATATTATTACTGTGATTAACCAAGGCGGTTTACACA 1056
QY 394 GluAsnThrAsn-----PheGlyAsnArgLeu 402
DB 1057 AGAAATTTTCGAGACATTTGATCTTATGATATTTTCTTTTACAGGTAAACCGATG 1116
QY 403 ValGlyIleSerAsnArgAspAlaProThrTyrSerAsnThrIleThrGluThrLeuTyr 422
DB 1117 -----GCTTTACACATCTAATGATGATCGCAACATATCTGG 1155
QY 423 GlyGluArgThrGly-----SerProThrThrLysThrIleArgProPheGluSer 439
DB 1156 GGAAGCGGTCATGACATATTAATTTCTCAAGACATCAACGATATTTCTTTTATAGA 1215
QY 440 Tyr-----LysValSerIleValThrAspArgGlnSerProValSerPro 455
DB 1216 AACAAACCTATGATAGATCGAAATTTGACACATAGACATGACTACATCAATATA--- 1272
QY 456 IleGlnProHisPheIleIleAsnGlnIleGluLeuTyrLeuAsnGlySerSerAsnAsn 475
DB 1273 -----TATGAAAGATATTTTTCGATATGACAGTAA 1305
QY 476 ThrLeuLysTyrSerAlaGlyLysLeu-----SerAsnTyrGlnAsnThrThrPhePhe 494
DB 1306 GATATTCGATATTCATCCAAATTCACAAATAGAAATATATTAAGAACTGATCTTAT 1365
QY 495 GlnPheProArgLysLysAspCysAsnLeuValIleAspProGlyCysSerProAsnPhe 514
DB 1366 ATGATTTCCAAACAAACATGAAATAAT----- 1392
QY 515 AsnAsnTyrSerHisIleLeuSerHisPheSerLeuPheThrTyrSerTyrValIleGly 534
DB 1393 AAAGATATGCTCATATCTTATGATATATAAACTGATATATATATTTCAAGATTT 1452
QY 535 LeuGlnLeuGlnIleLeuAspThrGlyValLeuGlyTyrThrHisSerSerValAspArg 554
DB 1453 AGAGAAAGAAAGAGTT-----GCATTTAGTTGACACATATGATGTTGATTTTC 1503
QY 555 TyrAsnAlaIleSerAspLysIleIleThrMetIleProAlaIleLysGlyAsnAsnLeu 574
DB 1504 CAATATACAAATGATTAATGATTAACATCACCCAAATTCACAGCTCTTAAGAGTAA 1563
QY 575 AspThrAsnSerLysValIleGluGlyProGlyHisThrGlyGlyAsnLeuValTyrLeu 594
DB 1564 AGTTCTGATTTCAAAATTTGTGAAGTCTGTGTCAACAGGTCGAGACTGTGTAATTTCT 1623
QY 595 GlnSerGlnGlyArgLeuGluIleThrCysGluThrProAsnSerThrGlnSerTyrPhe 614
DB 1624 AAAGATAGTATGATTTTATGATTAATTT---TTAAAAAATGTTTCTGACAAATTCAA 1680
QY 615 IleArgLysArgTyrAlaThrAsnGlyAlaGlyAsnThrLeuProAsnIleSerLeuThr 634
DB 1681 GTACGATATCGTTATGCTACTATGCT-----CCAAAGCAACACAGATATTC 1725
QY 635 IleProGlyValIleGlyIleProProGlnArgLeuAsnAsnThrPheSerGlyThrAsn 654
DB 1726 TTAAACCGGAATGATATCTAATAGTGTGAG---CTCCTAGTACCACTTCCGCCCAAAAC 1782
QY 655 TyrAsn-----AsnLeuGlnTyrGlyAspPheGlyTyrPheGlnPheProSerThrVal 672

Db 1763 CCAATGCTACGATTTTACATATGCAATTTTGCATATGTACATTTCCAAAGAACGTT 1842
Qy 673 -----ThiLeuProLeuAsnArgAsnIleProPheIle 683
Db 1843 CCAATTAACATTTGAAGAGAGAGACACTTATTAATGACC-----TTA 1887
Qy 664 PheAsnArgAlaAspValSerAsnSerIleLeuIleIleAspValIleGluPheIlePro 703
Db 1888 TATGTATACCAAAATCATTTATATATATA---TATATGTACAAATGCAATTTATTCOA 1944
Qy 704 IleThrSerSerMetHisGlnAsnArgGluLeuGlnLeuGluThrIleGlnThrIle 723
Db 1945 ATCACTCAATCTGTATATAGATTTATACAGAGAACAAATATATGAAAAACACGAAAAATA 2004
Qy 724 IleAsnThrPhePheThrAsnHisThrIleThrLeu 735
Db 2005 GTGAATGATTTATTTGTTATATTAACAAAGTTCTT 2040
RESULT 4
PCT-US94-07902-30
Sequence 30, Application PC/TUS9407902
GENERAL INFORMATION:
APPLICANT: APPLICANT: Street address: 4980 Carroll Canyon Road
APPLICANT: City: San Diego
APPLICANT: State/Province: California
APPLICANT: Country: US
APPLICANT: Postal code/Zip: 92121
APPLICANT: Phone number: (619) 453-8030 Fax number: (619) 453-6991
APPLICANT: Telex number:
TITLE OF INVENTION: Materials and Methods for the Control of
TITLE OF INVENTION: Calliphoridae Pests
NUMBER OF SEQUENCES: 31
CORRESPONDENCE ADDRESSES:
ADDRESSES: David R. Saliwanchik
STREET: 2421 N.W. 41st Street, Suite A-1
CITY: Gainesville
STATE: FL
COUNTRY: USA
ZIP: 32606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US94/07902
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Saliwanchik, David R.
REGISTRATION NUMBER: 31,794
REFERENCE/DOCKET NUMBER: MA79
TELECOMMUNICATION INFORMATION:
TELEPHONE: 904-375-8100
TELEFAX: 904-372-5800
INFORMATION FOR SEQ ID NO: 30:
SEQUENCE CHARACTERISTICS:
LENGTH: 2061 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
PCT-US94-07902-30
Alignment Scores:
Pred. No.: 1.61e-85 Length: 2061
Score: 899.50 Matches: 247
Percent Similarity: 48.3% Conservative: 126
Best Local Similarity: 32.0% Mismatches: 270
Query Match: 23.2% Indels: 129
DB: 6 Gaps: 30

US-10-783-417-2 (1-735) x PCT-US94-07902-30 (1-2061)
Qy 1 MetAsnGlnAsnAspAsnAsnGluTyrGluIleIleAspSerHisThrSerProTyr 20
Db 1 ATGAATCATATCAAAATAGAAATGAATATGAAATATTCATATGCTCCATCCAAATGGTTT 60
Qy 21 PheProAsnArgAsnSerAsnAspSerArgTyrProTyrThrAsnAsnProAsnGlnPro 40
Db 61 -----ACCAAGTCTAAATACATATCTGATATTCATATTCATAGCAAAATMAACCAATACCA 114
Qy 41 LeuGlnAsnThrAsnTyrLeuGluTyrLeuAsnMetCysGlnGlyAsnThrGlnTyrGly 60
Db 115 CTGAATAACACAGCAATTCACAAAGATTTGGCTCAATGCTGTCAAGATATCAACAAATATGGC 174
Qy 61 AspAsnProGluThrPheAlaSerAlaAspThrIleAlaIleValSerAlaGlyThrIle 80
Db 175 AATAATGGGGGAATTTTGTAGTTCTGAATCTATGTTGAGTTAGTGAAGATTTATTT 234
Qy 81 ValSerGlyThrLeuLeuAlaGlyIleGlyGlyLeuThrSerIleSerGlyProIleGly 100
Db 235 GTAGTAGGAACTATGTTAGCA-----GCTTTTGCTGCCCT----- 270
Qy 101 IleIleGlyAlaIleIleIleSerPheGlyThrLeuIleThrValPheThrProAlaGly 120
Db 271 GTCTAGCTGACGATATATATCTTTGGACCTTGTTGGATCTTTGG-----CAAGCA 327
Qy 121 GluGlnAspLysThrValTyrThrGlnPheIleLysMetGlyGluIlePheValAspThr 140
Db 328 TGTGACCTGCAAAATGTTTGGACAGATTTGTTAAACCGGA-----GGAAG 375
Qy 141 ProLeuThrGlu-----SerIleLysGlnLeuLysLeuGlnThrLeuGluGlyPhe 157
Db 376 CCTATACAGAAATATGATTAACAACTAATATGATTAATGCTTATCTGTAACACTATA 435
Qy 158 ArgGlnIleLeuGlnSerTyrAsnThrAlaLeuAspAspTyrPargLysLeuLysArgLeu 177
Db 436 AAAAATCAACTGATTAATATCAAGAAATTTTCGATAAATGGAGCCAGCAGC----- 489
Qy 178 GlnAlaProGlyLeuProProSerSerAlaLeuGlnGlnAlaAlaLeuThrLeuLysIle 197
Db 490 -----ACACAGCTAATGCTAAGCAGTACAGATCTCTTACT 528
Qy 198 ArgPheGluAsnValHisAsn-----AspPheIleArgGluIleProGlyPheGln 214
Db 529 ACCTTAGAACCTATATAGATTAAGATTAAGATTAAGATTAAGATTAAGATTAAGATTAAG 588
Qy 215 LeuGluThrTyrLysThrLeuLeuProIleTyrAlaGlnAlaAlaAsnPheHisLeu 234
Db 589 ATACCAACA-----CTCCCTGATATGCAAAATAGCTACTTGGACCTTG 633
Qy 235 AsnLeuLeuGlnGlnGlyAlaGluLeuAlaAspGluTyrAsnAlaAspIleHisProSer 254
Db 634 AATTTATTAACATGCTGCTACCTATTCATATATATG-----CTGCAAAATCA 684
Qy 255 GlnIleGluProAsnAlaGlyThrSerAspAspTyrTyrLys---LeuLeuLysGluAsn 273
Db 685 GGTATTAATCCAGATCTTCATCAATCATATATCAATCAAGGGCTATTTAAACGTAAA 744
Qy 274 IleProLysTyrSerAsnTyrCysAlaAsnThrTyrArgThrGlyLeuLysAsnLeuArg 293
Db 745 ATACAGAAATATATGACTATTTGTATACAAACGTCACATGACGCTACTATGATGTAGA 804
Qy 294 AspGluProAsnMetLysTyrSerIlePheAsnAspTyrArgArgTyrMetThrIleThr 313
Db 805 ACTAATATTAACGCAACATGCAATATATATATATCTTACGTTTGAATATACCTAATCT 864
Qy 314 ValLeuAspThrIleSerGlnPheSerLeuTyrAspIleLysArgTyrArgAspSerIle 333
Db 865 GTGTATAGATCTTATTTGCTATTTTCCAAATATATGACCAAGAAAAATAT----- 912
Qy 334 GlyGlyIleGluValLysGlyIleLysAsnGluLeuThrArgGluIleTyrThrGlu 353
Db 913 -----CCAATGAGATTAAATCTGAATCTTACAGAGAGATTAT---ACGAAT 957

QY	354	lleasnphesapargleuproglnleuarkyalglnproasleuula1lethmeGlutyr	373
Db	958	GTAAATTCAGATACATT-----AGAACCAATPAACAAGAACTGGAAAT	999
QY	374	AsnleuthrAglAserPhelyleuSerpheleugGlnPhe1lePheTythr	393
Db	1000	GGATTAACTAA---AATCCTACATTAATTACTGGATAAACCAAGGCCGTTTACCA	1056
QY	394	GluAenthraen-----PheGlyAsnargleu	402
Db	1057	AGAATTCGAGACATCTTGATCCTATGATATTTTCTTTACAGAGTAACAAGATG	1116
QY	403	ValGly1leSerAsnargAspAlaProthrytyserAenthrllethrglnThleuty	422
Db	1117	-----GCCCTTAACAATCATATGATGATGATCGACATTAATTCG	1155
QY	423	GlyGluArgthrgly-----SerProthrthlystrhileargProPheluser	439
Db	1156	GGAGCGGTCATGACATATTAATTTCTCAAGACACATCAAAAGTAATTCCTTTTATGA	1215
QY	440	Tyr-----LysValser1leValThrasparginserProProValSerPro	455
Db	1216	AACAACCTATGATTAAGTGTGAATTCGACACTAAGAGTACTCAGATATAATA---	1272
QY	456	lleglnProhisphelle1leasnGln1leGluLeutyryleuasnGlySerSerAsn	475
Db	1273	-----TATGAAATGATATTTTTCGAAATGACAGCTGA	1305
QY	476	ThreuleutyryserAlaGlyGlySerleu---SerAnthyrglnAsnthrThrophe	494
Db	1306	GTATTTGCATATTCATCCAAATTCACAAATTAAGAAATTAATTAATAAGACATGATCTTAT	1365
QY	495	GlnPheProArglylyAspCysAsnleuVal1leAspProGlyCysSerProAsn	514
Db	1366	ATGATTCCAAAACAACATGGAAT-----1392	1392
QY	515	AsnAnthyryserh1leleuSerh1sphereserleupherthytyserTyVal1leGly	534
Db	1393	AAAGATATGTGCTACTCTATCGATATAAACTGATAATTAATTAATTTTCAGTACTT	1452
QY	535	LeuGlnleuGln1leleuAspThrglyValleuGlyTPThrh1serSerValAspArg	554
Db	1453	AGAGAAAGAAAGAGATT-----GCATTTAGTTGACACATACATGAGTTGATTTTC	1503
QY	555	TyrAsnAla1leSerAspLyse1lelethrhethrleProAla1lelyGlyAsnAsnleu	574
Db	1504	CAAAATACATATGATTTAGATPAACATCCCAAAATCCACGCTCTAAAGCTTTGAAGTA	1563
QY	575	AspThrasenseTyVal1leGlnGlyProGlyh1sthrGlyGlyAsnleuValTyxleu	594
Db	1564	AGTTCTGATTCGAAATATGTGAAGAGTCCTGCTCAACAGGTGAGACTTGTAATCTT	1623
QY	595	GlnSerglnGlyArgleuGln1leThrcysGlnThrProAsnSerThrglnSertyrPhe	614
Db	1624	AAAGATAGTATGATTTTAAAGATTAGATT---TTAAAAAATGTTTTCGCAAAATATCAA	1680
QY	615	l1leargleuArgTyVal1leThrasnGly1leGlyAsnthrleuProAsn1leSerleuTh	634
Db	1681	GTACGATATCGTATGTACTTAATGCT-----CCMAAGACAAGATATTC	1725
QY	635	l1leProGlyVal1leGly1leProProGlnArgleuAsnAenthrrPheSerglyThrasn	654
Db	1726	TTAACCGAATAGATCTATTAAGTGTGAG---CTCCCTAATACCACTTCCCGCAAAAC	1782
QY	655	TyrAsn-----AsnleuGlnTytylyAspPheGlytyrPheGlnPheProSerThryAl	672
Db	1783	CCAATGTCTACAGATTTTAACATATGACATTTTGGATATGTAACATTTCCAAAGAACAGTT	1842
QY	673	-----ThreuproleuAsnArgAsn1leProPhe1le	683
Db	1843	CCAAATAAACATTTGAAGAGAAACCTTTATTAATGAACC-----TTA	1887

```

Oy 684 pheasmnrglaasrValSerAsnSerlleleuilelleapLysilleGluPellePro 703
Db 1888 TATGTGACACCAAAATCTCTATTAATAATA--TATATTGACAAATCGAATTTATTCCA 1944
Oy 704 lIethserseemethiagInasmarGluLysGlnLysleuGluThrIleGlnThrLys 723
Db 1945 ATGACTGATCTGTATTAGATTATACAGAGAGCAAAATATTAAGAAAAAACAAGAAATA 2004
Oy 724 lIeAnthrhrpethrAenhiSthrLysThrleu 735
Db 2005 GTGAATGATTTATTTGTTAAATTAACAAAGTTCTT 2040

RESULT 5
US-08-315-468-3
; Sequence 3, Application US/08315468
; Patent No. 5554534
; GENERAL INFORMATION:
; APPLICANT: Michaels, Tracy Ellis
; APPLICANT: Foncecrada, Luis
; APPLICANT: Narva, Kenneth E.
; TITLE OF INVENTION: Process for Controlling Scarab Pests
; TITLE OF INVENTION: with Bacillus thuringiensis Isolates
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: David R. Saliwanchik
; STREET: 2421 N.W. 41st Street, Suite A-1
; CITY: Gainesville
; STATE: FL
; COUNTRY: USA
; ZIP: 32606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/315,468
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/014,941
; FILING DATE: 01 FEB 1993
; APPLICATION NUMBER: 07/828,430
; FILING DATE: 30-JAN-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/808,316
; FILING DATE: 16-DEC-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Saliwanchik, David R.
; REGISTRATION NUMBER: 31,794
; REFERENCE/DOCKET NUMBER: MA73, C2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 904-375-8100
; TELEFAX: 904-372-5800
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3507 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHEetical: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Bacillus thuringiensis
; STRAIN: kumamotoensis
; INDIVIDUAL ISOLATE: 50C
; IMMEDIATE SOURCE:
; LIBRARY: LambdageM-11(tm) library of L. Foncecrada
; CLONE: 50C(b)
US-08-315-468-3

```

Pred. No.: 6,55e-72 Length: 3507
 Score: 775.50 Matches: 217
 Percent Similarity: 46.44 Conservative: 136
 Best Local Similarity: 28.54 Mismatches: 289
 Query Match: 20.04 Indels: 119
 DB: 2 Gaps: 27

US-10-783-417-2 (1-735) x US-08-315-468-3 (1-3507)

QY 5 AenAaPaenAenGluTyrGluIleIleAaPserHisThrSerProTyrPheProAenArg 24
 DB 10 AATATTCAAATATGATATATGAAATTTATAGTCGACACCTTCT-----ACATCT 57
 QY 25 AenSerAaenAaPser---ArgTyrProTyrThrAaenAaenProAaenGlnProAenGlnAa 43
 DB 58 GATACCAATGATTTCAACAGATACCTTTTGGCAATGAGCCAAACAAATCCGCTACAAAT 117
 QY 44 ThrAenTyrLeuGluTyrLeuAaenMetCysGlnGlnAaen---ThrGlnTyrGlnAaen 62
 DB 118 ATGATTTTAAAGATTTATTTAAATGTCGCGGAAAGTTAGTGAATACCTCGTTCA 177
 QY 63 PheGluThrPheAaSerAaAaPthrIleAaIleAaAaSerAaGlyThrIleAaAaSer 82
 DB 178 CCTGAGGATTTCTTAAAGCAGCAAGAT-----GCAGTTAAGCGCGCAATGATATAGTA 231
 QY 83 GlyThrLeuLeuAaGlyIleGlyLeuThrSerIleSerGlyProIleGlyIleIle 102
 DB 232 GGTAATATCTAAACAGGTTTAGCG-----CTTCATTTGTT 267
 QY 103 GlyAlaIleIleIleSerPheGlyThrLeuIleThrValPheTyrProAaGlyGln 122
 DB 268 GGGCGGATAGTACGCTTATATCTCAACTATGATATCTGTCGCTTCA-----AAA 321
 QY 123 AaPlyThrValTyrThrGlnPheIleLeuMetGlyGlnIlePheValAaPthrProAen 142
 DB 322 CAAGAAGATCAATGGCAATTTTATGACAAAGTAAGAACTCATTAATCAAAAAATTA 381
 QY 143 ThrGluSerIleLeuGlnLeuLeuLeuGlnThrLeuGlnGlyPheArgGlnIleLeuGln 162
 DB 382 GCGAATATGCAAGAAATTAAGCCGCTTTCGAATTCGAAGGCGCTAGGAAATTAATACGA 441
 QY 163 SerTyrAenThrAaLeuAaAaPserTyrArgIleLeuLeuLeuGlnAaAaProGlyLeu 182
 DB 442 TTATATCTTAACTCGCTTGAAGAGTGAAGAA----- 474
 QY 183 ProProSerSerAaLeuGlnGlnIleAaAaLeuThrLeuLeuIleAaPheGluAaVal 202
 DB 475 AATCCAAATGGTTCA-----AGAGCTTACGAGATGTTGAAATCGAATTTGAAATCCG 528
 QY 203 HisAaAaPserPheIleArgGluIleProGlyPheGlnLeuGluThrTyrIleThrLeuLeu 222
 DB 529 GATAGTTATTTACGCAATATATGCAATCTTTTCAGTCGACAAATTTTGAATACCAATTC 588
 QY 223 LeuProIleTyrAlaGlnAaAaAaenPheHisLeuAaenLeuGlnGlnGlyAlaGln 242
 DB 589 CTTACAGATATATACATGCGACCAAACTTACTTTATTAAGGACCCATCAATTC 648
 QY 243 LeuAaAaPserGluTyrAaenAaAaPserIleHisProSerGlnIleGluProAaenAaGlyThr 262
 DB 649 TTGGAGAAAGATGGGGATTG-----CTTCAAGCACT 681
 QY 263 SerAaAaPserTyrTyrLeuLeuLeuLeuLeuAaenIleProLeuTyrSerAaenTyrCysAa 282
 DB 682 ATTAATTAATCTACTATATATGCTCAATGAAACTTACGCAATATTTCTGACCACTGTGTA 741
 QY 283 AaenThrTyrArgThrGlyLeuLeuLeuAaenAaAaPserGluProAaenMetLeuTyrPserIle 302
 DB 742 AAGGTATGAAACCTGGTTAGCAAAATTAAGGCTGACGCGCTAAACATGCAATGATGAC 801
 QY 303 PheAaenAaPserTyrArgArgTyrMetThrIleThrValLeuAaPthrIleSerGlnPheSer 322
 DB 802 TATAACCAATTCCTGAGAGAAATGACATTCGAGGTGTAGACGTTGTTGCAATTAATTTCA 861

QY 323 LeuTyrAaPserIleLeuAaArgTyrArgAaPserIleGlyGlyIleGluValIleGlyIleLeu 342
 DB 862 AACTATGATAGCCGTACGTAT-----CCATCGGCAACAACA 897
 QY 343 AaenGluLeuThrArgGluIleTyrThrThrGluIleAaenPheAaPserLeuProGlnLeu 362
 DB 898 GCTCAGCTTACAGGAGGATATATACATTCACATTCGCGCGGTAGATGTGCTTAATAT 957
 QY 363 ArgValGlnProAaenLeuAaAaThrMetGluTyrAaenLeuThrArgAaAaSerPheLeu--- 381
 DB 958 GGCTCCTGATAGCAAAAGCACCTCTTTCAGAAATAGAAAGACCGCTATTCGTCA 1017
 QY 382 -----LeuPheSerPheLeuGlnGlnPheIlePheTyrThrGluAaenThrAaPhe--- 398
 DB 1018 CCTCATGTGTTGATATATATACAGGACTCACAGTTTATACAAAACGATGCTTCACT 1077
 QY 399 -----GlyAaenAaGluValGlyIleSerAaenAaPser 409
 DB 1078 TCTGATCGTTATATGATATATGCGCTGCTCATCA-----ATAAGCTATTAAGCAT 1128
 QY 410 AlaProThrTyrSerAaenThrIleThrGluThrLeuTyrGlyGluArgThrGlySerPro 429
 DB 1129 ATCGGTACG---AGTATGACCTTTACACAG---ATGATGAAACCAATCAAAATTTACAA 1182
 QY 430 ThrThrIleThrIleArgProPheGluSerTyrIleVal----- 442
 DB 1183 AGTACTACCAATTTT---GATTTTACGAATTAACGATATTTACAGACTTATCAATGAT 1239
 QY 443 SerIleValThrAaPserGlnSerProProAaAaSer-----ProIle 456
 DB 1240 GCAGTACCTCTGATATATGTTTACCTGCTTATACGATATACATTTTGGAAATGCGACAA 1299
 QY 457 GlnProHisPheIleIleAaenGlnIleGluLeuTyrLeuAaenGlySerSerAaenThr 476
 DB 1300 ACCGAGTTTATATGTAATCAATTC-----AATATATCC 1335
 QY 477 LeuTyrTyrSerAaAaGlySerLeuSerAaenTyrGlnAaenThrPhePheGlnPhe 496
 DB 1336 AGAAG-----ACGTTAAAGTATTAACCACTTCCAAAGATATTTATA 1377
 QY 497 ProArgGlyLeuAaPserCysAaenLeuValIleAaPProGlyCysSer-----ProAaPhe 514
 DB 1378 GATCGACAAAGATTCGCAATTAAGATTCCTCCAGAACTTCAGGTCAACCAATATAC 1437
 QY 515 AaenAaenTyrSerHisIleLeuSerHis-----PheSerLeuPheThrTyrSer 530
 DB 1438 CAGTCATATACCAATATATAGTATATATATATATATATATATATATATATATATATAT 1497
 QY 531 TyrValIleGlyLeuGlnLeuGlnIleLeuAaPthrGlyValLeuGlyTyrThrHisSer 550
 DB 1498 TATGTA-----CCGTATTTCTTGACACATCCG 1527
 QY 551 SerValAaPserTyrAaenAa 570
 DB 1528 AGTGCAGATCTTAAACAAATACAGTTTAAAGTGGCGAAATCACCAATACACAGGGGCGAAG 1587
 QY 571 GlyAaenAaenAaPserThrAaenSerLeuValIleGlnGlyProGlyHisIleThrGlyGlyAaen 590
 DB 1588 TCTACACACATAGCGCAAAATCTTATATATATATATATATATATATATATATATATATAT 1647
 QY 591 LeuValTyrLeuGlnSerGln---GlyArgLeuGlnIleThrCysGluThrProAaenSer 609
 DB 1648 TTAGTGCGTTTAAAGCAAGCGCATGGAAGTTGAGTTTACAGATGATCTTCCAGAGCTC 1707
 QY 610 ThrGlnSerTyrPheIleArgLeuArgTyrAlaAaAaenGlyAlaAaAaenThrLeuPro 629
 DB 1708 ---CAACATTCCTGATTCGGAATTCGTTACGCTTCAATGAACTAGT----- 1752
 QY 630 AaenIleSerLeuThrIleProGlyValIleGlyIleProProGlnArgLeuAaenAaenThr 649
 DB 1753 -----TAT 1806
 QY 650 PheSerGlyThrAaenTyrAaenAaenGlnTyrGlyAaPheGlyTyrPheGlnPhePro 669

```

Db      1807 TATCTTAATATAAAATGAAATGATTACATATATATATGTAATATATATGAAATATCCA 1866
Qy      670 SerThraValThrLeuProLeuAsnArgAsnIleProPheIlePheAsnArgAlaAspVal 689
Db      1867 AGAGTCATTTACGATTAATGCTTCTTCAAACATACAGAGCTTATCTATATAGATACAAAGC 1926
Qy      690 SerAsnSerIleLeuIleIleAspLysIleIleGluPheIleProIleIleIleSerSerMetHis 709
Db      1927 AATCAAAATTTATTTATTTATTTAGACCGAATCGAATTCATCCAGATGATGACATATGAA 1986
Qy      710 GluAsnArgGluLysGlnLysLeuGluThrIleGlnThrLysIleAsnThrPhePheThr 729
Db      1987 GCGGAACGCGAT-----TTAGAACGCGCAAAAGAAAGCATGTAATGCTTTTACG 2037
Qy      730 Asn 730
Db      2038 AAT 2040

RESULT 6
US-08-315-468-5
Sequence 5, Application US/08315468
Patent No. 5554534
GENERAL INFORMATION:
APPLICANT: Michaels, Tracy Ellis
APPLICANT: Foncerada, Luis
APPLICANT: Nayra, Kenneth E.
TITLE OF INVENTION: Process for Controlling Scarab Pests
TITLE OF INVENTION: with Bacillus thuringiensis Isolates
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESSES:
ADDRESSER: David R. Saliwanchik
STREET: 2421 N.W. 41st Street, Suite A-1
CITY: Gainesville
STATE: FL
COUNTRY: USA
ZIP: 32606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/315.468
FILING DATE:
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/014.941
FILING DATE: 01 FEB 1993
APPLICATION NUMBER: 07/828.430
FILING DATE: 30-JAN-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/808.316
FILING DATE: 16-DEC-1991
ATTORNEY/AGENT INFORMATION:
NAME: Saliwanchik, David R.
REGISTRATION NUMBER: 31,794
REFERENCE/DOCKET NUMBER: MA73.C2
TELECOMMUNICATION INFORMATION:
TELEPHONE: 904-375-8100
TELEFAX: 904-372-5800
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 1953 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Bacillus thuringiensis
STRAIN: tolworth1

```

```

; INDIVIDUAL ISOLATE: 43F
; IMMEDIATE SOURCE:
; CLONE: E. coli XL1-Blue (pML798-4), NRRL B-18291
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..1953
US-08-315-468-5

Alignment Scores:
Pred. No.: 1,056-61 Length: 1953
Score: 676.00 Matches: 209
Percent Similarity: 43.84 Conservative: 117
Best Local Similarity: 28.1% Mismatches: 283
Query Match: 17.5% Indels: 136
DB: Gaps: 29

US-10-783-417-2 (1-735) x US-08-315-468-5 (1-1953)

QY 1 MetAenGlnaAnaenAAspAaenGlnuTyrguIlelleapSerHieRhpProTyR 20
DB 1 ATGATATCCAAACAAAT--CGAAGTAATATATGATATGATTAAGCTTACCTTACAGCTGAA 57
QY 21 PheProAenAArgASeRAspAaSerAqTyRProTyRThcAsnaAnProAenGlnPro 40
DB 58 TTGCCA-----ACTTACCATATCATATCTCTTACGTACATCATCAATTCGACA 108
QY 41 LeuGlnAenThraenTyRlyeGluTrpLeuAenMetCyseGlnGlyAsnThrglnTyRgly 60
DB 109 CTAGAGAAATTAATTAATTAAGAAATTTTAAAGATGATCGACAGCAATTCACG----- 162
QY 61 AsphenPheGluThrPheAlaSerAlaapThrIleAlaIValaSerAlaGlyThrIle 80
DB 163 -----GAAGTCCTAGACAGCTCTCTACAGTAAAGATGCGATTGGACAGGAATTTCT 213
QY 81 ValSerGlyThrLeuLeuAlaGlyIleGlyGlyLeuThrSerIleSerGlyProIleGly 100
DB 214 GTTGTAGACAGATTTTGGCTGTGTAGAG-----GTTCCA 249
QY 101 IleIleGlyAlaIleIleIleSerPheGlyThrLeuIleThraValPheTrpProAlaGly 120
DB 250 TTGCTGGGGGGGCTCATCTTCAATTTATCAATCATTTCTTAAAGCTATAGGCCAAGT--- 306
QY 121 GluGlnAapLystrHValTrpThrGlnPheIleLyseMetGlyGluIlePheValaapThr 140
DB 307 -----GATGCTGACCCATGAGAGGCTTTTAAAGGACAAGTGAAGTACGTATGATAG 360
QY 141 ProLeuThrgInuSerIleLyseGlnLeuLyseGlnInThrLeuGlnGlyPheArgGlnIle 160
DB 361 AAAATATAGAGAGATATGCTTAAAGTAAAGTAAAGCTCTTGACAGATACAGGGCTTCAAAATAT 420
QY 161 LeuGlnSerTyraenThraLeuAaSpAapTrpArgLyseuLyseArgLeuGlnAlaPro 180
DB 421 TTGTGAAGTTATGTAAATGCGTGTGATTCCTCGAAGAA----- 459
QY 181 GlyLeuProProSerSerAlaLeuGlnGlnAlaAlaLeuThrLeuLyseIleArgPheGlu 200
DB 460 ---GCGCCTGTAAATTTTACGAAGTGAAGAACCCAAAGATCGATATAGAAACAATTTTCT 516
QY 201 AsnValIleAaenAapPheIleArgGluIleProGlyPheGlnLeuGluThrTyRlystr 220
DB 517 CAAGCAGAAAGCATTTTGTATATTCACCGCTCATTTTGGCGTTTCAATTCGAAGTT 576
QY 221 LeuLeuLeuProIleTyRAlaGlnAlaAlaAenPheHieLeuAaenLeuLeuGlnGlnGly 240
DB 577 CTGTTTTCACCAACATATGCGACAACTCCAAATATACATTTATGCTATTTAAAGATGCT 636
QY 241 AlaGlnLeuAlaapGlyTrpAsnAlaapIleHisProSerGlnIleGluProAaenAla 260
DB 637 CAAGTTTGTGAGAAATATGGGCAAT----- 663
QY 261 GlyThrSerAaPAP-----TyTrTyRlyseLeuLeuLyseGluAaenIleProLystr 277
DB 664 ---TCTTGAAGATATGCTGAAATTTTATCAAGACATTAATAATTATTCGACCAATATAC 720

```



```

/ ORIGINAL SOURCE:
/ ORGANISM: Bacillus thuringiensis
/ STRAIN: colworthi
/ INDIVIDUAL ISOLATE: 43f
/ IMMEDIATE SOURCE:
/ CLONE: E. coli XL1-Blue (pM1, 98-4), NRRL B-18291
/ FEATURE:
/ NAME/KEY: CDS
/ LOCATION: 1..1953
US-07-941-650A-3

Alignment Scores:
Pred. No.: 1,05e-61 Length: 1953
Score: 676.00 Matches: 209
Percent Similarity: 43.8% Conservative: 117
Best Local Similarity: 28.1% Mismatches: 283
Query Match: 17.5% Indels: 136
DB: Gaps: 29

US-10-783-417-2 (1-735) x US-07-941-650A-3 (1-1953)
QY 1 MetAenGlnAsnAsnAsnAsnGluTyrGluIleIleAspSerHisThrSerProTyr 20
DB 1 ATGAATCCAAACAAAT---CGAAGTGAATATGATACGATTAAGTTACACCTTAACACTGA 57
QY 21 PheProAsnArgAsnSerAsnAspSerArgTyrProTyrThrAsnAsnProAsnGlnPro 40
DB 58 TTGGCC-----ACTAACCATATATCAATATCTTTAGCTGACATCCAAATTCGACA 108
QY 41 LeuGlnAsnThrAsnTyrIleGluTyrPheAsnMetCysGlnGlyAsnThrGlnTyrGly 60
DB 109 CTAGAAGATTAATATTAAGAATTTTAAAGATACGACAGACAATTCGCG----- 162
QY 61 AspaAsnGluThrPheAlaSerAlaAspThrIleAlaAlaValaSerAlaGlyThrIle 80
DB 163 -----GAAGTCTAGACAGCTCTCACTAAGTAAAGATGAGTGGACAGGAATTTCT 213
QY 81 ValSerGlyThrLeuLeuAlaGlyIleGlyGlyLeuThrSerIleSerGlyProIleGly 100
DB 214 GTTGTAGGACAGATTTAGGCTGTAGG-----GTTCCTCA 249
QY 101 IleIleGlyAlaIleIleIleSerPheGlyThrLeuIleThrValPheThrProAlaGly 120
DB 250 TTTCGTGGGGCCCTCATCTTATTCATATTCATCTTTCTTAACGCTATATGCGCAAT--- 306
QY 121 GlnGlnAspLysThrValTyrThrGlnPheIleLysMetGlyGluIlePheValAspThr 140
DB 307 -----GATGCTGACCCATGGAAGGCTTTATGCGACAGTGAAGTACTGATAGTAAG 360
QY 141 ProLeuThrGlnSerIleLysGlnLeuLysLeuGlnThrLeuGlnGlyPheArgGlnIle 160
DB 361 AAAATAGAGAGATGCTAAAGTAAAGCTTTGACAGACTTACAGGGCTTCCAAATTAAT 420
QY 161 LeuGlnSerTyrAsnThrAlaLeuAspAspTyrPargLysLeuLysArgLeuGlnAlaPro 180
DB 421 TTTCGAAGATTAATGTAATGCGTTGATCTCTGGAAGAA----- 459
QY 181 GlyLeuProProSerSerAlaLeuGlnGlnAlaAlaLeuThrLeuLysValLeuArgPheGln 200
DB 460 ---GGCGCTGTAATTTAGCAAGTGAAGAGCCAGATGCAATAGAGAACTTTTCTCT 516
QY 201 AsnValHisAsnAspPheIleArgGluIleProGlyPheGlnLeuGluThrTyrLysThr 220
DB 517 CAAGCAGAAAGCCATTTTCGTAAATTCATGCGCTCATTTTCGCTTTCCAAATTCGAAGTT 576
QY 221 LeuLeuLeuProIleTyrAlaGlnAlaAlaAsnPheIleLeuAsnLeuLeuGlnGlnGly 240
DB 577 CTGTTTCTACCAACATATGCAACAGTGCAGAAATACCATTTATGTGATATTAAGATGCT 636
QY 241 AlaGluLeuAlaAspGluTyrAsnAlaAspIleHisProSerGlnIleGluProAsnAla 260
DB 637 CAAGTTTTCGAGAAAGATGCGGATAT----- 663

```

```

QY 261 GlyThrSerAspAsp-----TyrTyrLysLeuLeuLysGlnAsnIleProLysTyr 277
DB 664 ---TCTTGAAGATATATGCTGAATTTATCAAGACAACTTAACCTTACGCAACATAC 720
QY 278 SerAsnTyrCysAlaAsnThrTyrArgThrGlyLeuLysAsnLeuArgAspGluProAsn 297
DB 721 ACTGACCATGTGTCAATTTGATATATGTGATTAATTAAGTTTAAGAGTTCAACTTAT 780
QY 298 MetLysTyrSerIlePheAsnAspTyrArgArgTyrMetThrIleThrValLeuAspThr 317
DB 781 GATGATGGGTCAAAATTTAACCCTTTTTCGACAGAAATACATTAACGTATATTAAGTCTA 840
QY 318 IleSerGlnPheSerLeuTyrAspIleLysArgTyrArgAspSerIleGlyGlyIleGln 337
DB 841 ATTGATATATATCCATTTTATGATATGCGGTATATCTCA----- 879
QY 338 ValIleGlyIleLysAsnGluLeuThrArgGluIleTyrThrThrGluIleAsnPheAsp 357
DB 880 ---AAAGGCTTAAACACAGACTAACAGACATTTTACAGATCCAAAT---TTTACA 933
QY 358 ArgLeuProGlnLeuArgValGlnProAsnLeuAlaThrMetGluTyrAsnLeuThrArg 377
DB 934 CTCAAATGCTCTTCAAGATGACCACTTTTTCGATATGAAACCTCATTCGAAAA 993
QY 378 AlaSerPheLysLeuPheSerPheLeuGlnGlnPheIlePheTyrThrGluAsnThrAsn 397
DB 994 CCT-----CATTTATGATATATTTGCTGGATGTAATTCATACG----- 1035
QY 398 PheGlyAsnArgLeu---ValGlyIleSerAsnArgAspAlaProThrTyr----- 413
DB 1036 -----CGTCTTCGACCTGTGTACTCTCGAAAGAAATTCCTTCAATATATGTCGTCT 1086
QY 414 -----SerAsnThrIleThrGluThrLeuTyr 422
DB 1087 AATTATAGTAAACTAGACACTAGATAGAGTCAATGATACATCACTTCCCATTTAT 1146
QY 423 GlyIleArgThrGlySerProThrThrTyrThrIleArgProPheGlnSer----- 439
DB 1147 GGAGATTAATCATTTGAACTTATACAAAAGCTA-----AGCTTGAATGACAAAAGTT 1200
QY 440 TyrLysValSerIleValThrAspArgGlnSerProProValSerProIleGlnProHis 459
DB 1201 TATCAACATATGATTAATACAGACATAGCGGCTTTCCGATGCGCAAGATA-----TAT 1254
QY 460 PheIleIleAsnGlnIleGluLeu---TyrLeuAsnGlySerSerAsnAsnThrLeuLys 478
DB 1255 TTTCGTGTACGAAAGTTGATTTTATGCAATATGTATGATCAAAAATGAAACT----- 1308
QY 479 TyrSerAlaGlyLysSerLeuSerAsnTyrGlnAsnThrThrPheGlnPheProArg 498
DB 1309 -----AGTACACAACTATGATTTCAAAAAGTACATAGGCTATTTAGGT 1353
QY 499 LysLysAspCysAsnLeuValIleAspProGlyCysSer-----ProAsnPheAsnAsn 516
DB 1354 GCACAGATTCATATGACCAATTAACCAAGAAACACAGATGAACCACTTGAAAAAGCA 1413
QY 517 TyrSerHisIleLeuSerHisPheSerLeuPheThrTyrSerTyrValIleGlyLeuGln 536
DB 1414 TATAGTCAATGACCTTAATTCGACAGATGTTCTTA-----ATGCGAG 1455
QY 537 LeuGlnIleLeuAspThrGlyValLeuGlyTyrThrHisSerSerValAspArgTyrAsn 556
DB 1456 GACCGTCCGTGAACAATTCATTTTACTTGTGACATAGAGATGATGACTTTTATAT 1515
QY 557 AlaIleSerAspLysIleIleThrMetIleProAlaIleLysGlyAsnAsnLeuAspThr 576
DB 1516 ACAATTTGATGCGTGAATAATTAATCTCACTTCAGATGAGTAAAGCATATGCTTGTCTCA 1575
QY 577 AsnSerLysValIleGluGlyProGlyHisThrGlyGlyAsnLeuValTyrLeuGln--- 595
DB 1576 GCGGCTTCATTAATTAAGTTCAGATTCACAGAGAGAAATTTTCTATTCCTTAAGAA 1635
QY -----SerGlnGlyArgLeuGluIleThrCysGluThrProAsnSerThrGlnSer 612

```



```

Db      1636 TCTAGTAATGCAATGCTAAATTTAAAGTTACTTAAATTCAGCAGCCTTGTACACGA 1695
Qy      613 TyrPheIIeArGleuArGTYrAlaThraEngIyAlaGleuThleuProkAnIleSer 632
      1696 TATCCGCTAGAGAAATACGCTATGCTTCAAC-----ACTRACCTACGA 1737
Qy      633 LeuThrIleProGlyValIleGlyIleProProGlnArGleuAsnAsnThrPheSerGly 652
      1738 CTTTTCGTG-----CAAAATTCACCAATGATTTTCTGTGC 1773
Qy      653 ThrAenTYrAen-----AsnleuGlnTYrGlyAAspPheGlyTYr-----PheGln 667
      1774 ATCTACATTAAATAAATACTATGATAATAGAT--GGTGAATTAAACATATCAACATTGAT 1830
Qy      668 PheProSerThrValIleThleuProleuAsnArgAnIleProPheIlePheAsnArgAla 687
      1831 TTCCGAACTAGT-----AATTCTAATATGGGATTTCTGCGTACATCAAAAT 1875
Qy      688 Asp-----ValSerAsnSerIleleuIleIleAspIlyIle 699
      1876 GACCTTATATATAGAGACGAATCTTTCGTTCTAATGAAAAAATCTATATAGATAGATA 1935
Qy      700 GluPheIleProIle 704
      1936 GAATTATCCACGTA 1950
Db

RESULT 8
PCT-US92-00040-1
Sequence 1, Application PC/TUS9200040
GENERAL INFORMATION:
APPLICANT: Donovan, William P.
APPLICANT: Rupert, Mark J.
APPLICANT: Slaney, Annette C.
TITLE OF INVENTION: BACILLUS THURINGIENSIS CYTILIC (b) TOXIN
TITLE OF INVENTION: GENE AND PROTEIN TOXIC TO COLEOPTERAN INSECTS
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSES: Panitch Schwarze Jacobs & Nadel c/o A.S.
ADDRESS: Nadel
STREET: 1601 Market Street, 36th Floor
CITY: Philadelphia
STATE: Pennsylvania
COUNTRY: U.S.A.
ZIP: 19103
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US92/00040
FILING DATE: 19920103
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/649,562
FILING DATE: 31-JAN-1991
ATTORNEY/AGENT INFORMATION:
NAME: Egoif, Christopher
REGISTRATION NUMBER: 27633
TELECOMMUNICATION INFORMATION:
TELEPHONE: 215-757-1590
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 2430 base pairs
TYPE: NUCLEIC ACID
STRANDEDNESS: double
TOPOLOGY: circular
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 144..2099

```

```

PCT-US92-00040-1
US-10-783-417-2 (1-735) x PCT-US92-00040-1 (1-2430)

Alignment Scores:
Pred. No.: 3,24e-60 Length: 2430
Score: 663.50 Matches: 217
Percent Similarity: 41.9% Conservative: 115
Best Local Similarity: 27.4% Mismatches: 300
Query Match: 17.1% Indels: 161
DB: 6 Gaps: 31

Qy      1 MetAenGlnAenAsnAspAsnAsnGluTYrGluIleIleAspSerHisThSerProTYr 20
      144 ATGATCCAAACAAAT---CGAAGTGAACATGATACGATAAAGTTACCACTTACAGTGA 200
Qy      21 PheProAsnArgAenSerAsnAspSerArgTYrProTYrThraAsnProAsnGlnPro 40
      201 TTGCCA-----ACTRACCATATCATATATCTTTAGCTGACATCAATTCGACACA 251
Qy      41 LeuGlnAenThraenTYrGlyGluThreleuAsnMetCysGln---GlyAsnThrGlnTYr 59
      252 CTAGAAAGATTAATTAATTAAGAAATTTTAAGATGACTGAGACAGACAGTTTACGAAAGT 311
Qy      60 GlyAspAsnPheGluThrPheAlaSerAlaAspThrIleAla--AlaValSerAlaGly 78
      312 CTAGACAAAC-----TCTACAGTAAAGAGATGAGAGTGGACAGCA 350
Qy      79 ThrIleValSerGlyThrleuLeuAlaGlyIleGlyGlyLeuThSerIleSerGlyPro 98
      351 AATTCTGTGTAGAGCAGATTTTATGCTGTATGAA----- 386
Qy      99 IleGlyIleIleGlyAlaIleIleIleSerPheGlyThrleuIleThrValIlePro 118
      387 GTTCATTGCTGGGGCCTCACTTCATTATTAATCAATATTCTTGAACATATAGGCA 446
Qy      119 AlaGlyGluAspIleThrValIleThrGlnPheIleIleMetGlyIleIlePheVal 138
      447 AGT-----GATGCTGACCCCATGGAAGGCTTTATGAGCAACATGGAATGATCATGATA 497
Qy      139 AspThrProleuThrGluSerIleGlyGlnleuIleGlnThleuGlnIlePheArg 158
      498 GATAGAAAATGAGAGATATGCTAAAGTAAAGCTCTTGACAGACTTACAGGCTTTCAA 557
Qy      159 GlnIleleuGlnSerTYrAsnThrAlaIleuAspAspTYrArgIleIleuIleArgGln 178
      558 AATAATTTGAAAGATATATGCTTAATGCTTAATTCCTGAAGAAA----- 602
Qy      179 AlaProGlyLeuProProSerSerAlaIleuGlnGlnAlaIleuThleuIleArg 198
      603 -----ACACCTTAAATTGCGAAGTAAAGAACCCAAAGATCGAATAAGGAACTT 653
Qy      199 PheGluAenValHisAsnAspPheIleArgGluIleProGlyPheGlnleuGlnThrTYr 218
      654 TTTTCTCAAGCAAGAAAGCATTTTGTAAATTCACAGCCGTCATTTGCAAGTTCCAAATTC 713
Qy      219 LysThrleuLeuLeuProIleTYrAlaGlnAlaIleAsnPheHisLeuAsnleuGln 238
      714 GAAGTGTCTTTCTTCAACATATGACACAGCTGCAATATACATTTATTTGCTATTAATA 773
Qy      239 GlnGlyAlaGluLeuAlaAspGluThrPheAlaAspIleHisProSerGlnIleGluPro 258
      774 GATGCTCAAGTTTGTGAGAAAGATGGGATATTC----- 809
Qy      259 AsnAlaGlyThrSerAspAspTYrTYrIleIleuLeuGlyGluAenIleProIleTYrSer 278
      810 TCAGAAAGATGTCGATTTATCATATGACAAATTAACCTTACG---CAACAAATACACT 866
Qy      279 AsnTYrCysAlaAsnThrTYrArgThrGlyLeuIleuIleAsnleuArgAspGluProAsnMet 298
      867 GACCATTTGTCAATTTGATATATGTTGAATTAATGTTTAAGAGTTCAACTATATGAT 926
Qy      299 LysTrpSerIlePheAsnAspTYrArgArgTYrMetThrIleThrValIleuAspThrIle 318

```

```

Db      927 GCATGGGTCATAATTACCGTTTTCGACAGAAATGACTTAACTGTATTAGATCTTAATT 986
      319 SerGlnPheSerLeuTyrAspIleValArgTyrArgAspSerIleGlyIleGluVal 338
      987 GATCTTTCCTCCATTTTATGATGATGTCGTTACTCA-----1022
      339 LysGlyIleLeuAsnGluLeuThrArgGluIleTyrThrGluIleAsnPheAspArg 358
      1023 AAAAGGTGTTAAACAGAACTTACAGACATTTTTCGATCCAAATT---TTTCACTC 1079
      359 LeuProGlnLeuArgValGlnProAsnLeuAlaThrMetGluTyrAsnLeuThrArgAla 378
      1080 AATACCTCTCAGGAGATGAGACCACTTTTTCAGATAGAAAACCTATTCGAAAACCT 1139
      379 SerPheLeuLeuPheSerPheLeuGluGlnPheIlePheTyrThrGluAsnThrAsnPhe 398
      1140 -----CATTTATTTGATTTATTTACAGGGATTTGAAATTTTCATACG-----1178
      399 GlyAsnArgLeu---ValGlyIleSerAsnArgAspAlaProThrTyr-----413
      1179 -----CGTCTTCACACCTGGTATCTCGGAAAAGATTTTCAATTATGCTGTGTAAT 1232
      414 -----SerAsnThrIleThrGluThrLeuTyrGly 423
      1233 TATGTAGAAACTAGACCTAGTATAGATCTAGTAAGACATTAATCCCATTTTATGGA 1292
      424 GluArgThrGlySerProThrThrIleValArgProPheGluSer-----Tyr 440
      1293 GATTAATCTTACGACCTGATACAAAGTTA-----AGCTTGATGACAAAAAGTTTAT 1346
      441 LysValSerIleValThrAspArgGlnSerProProValSerProIleGlnProHisPhe 460
      1347 CGAATCTATGCTTAATACAGACGTAGCGCTTGCCCAATGCGCAAGTAA-----TATTTT 1400
      461 IleIleAsnGlnIleGluLeu-----TyrLeuAsnGlySerSerAsnAsn 475
      1401 GGTGTTCAGAAAGTTGATTTTACTCAATATGATGATCAAAAAAATGAACTAGTACACA 1460
      476 ThrLeuLysTyrSerIleGlySerLeuSerAsnTyrGlnAsnThrThrPhePheGln 495
      1461 ACATATGATTCAAAAAAGAAACAATGCGCATGTAGTGACACAGGATTTCTATGACCAATTA 1520
      496 PheProArgLysLysAspCysAsnLeuValIleAspProGlyCysSerProAsnPheAsn 515
      1521 CCAACCAAGAAACAACAT-----GAACTCACTTGAAAA 1553
      516 AsnTyrSerHisIleLeuSerHisPheSerLeuPheThrTyrSerTyrValIleGlyLeu 535
      1554 GCATATGATCATCAGCTTATTAATTCGCGGATGTTCTTA-----ATG 1595
      536 GlnLeuGlnIleLeuAspThrArgLysValLeuGlyTyrThrHisSerSerValAspArgTyr 555
      1596 CAGGACCGTCGTGGAACAATTCATTTTACTTGGACACATGAGAGTGAAGCTTTTAT 1655
      556 AsnAlaIleSerAspLysIleIleThrMetIleProAlaIleLysGlyAsnAsnLeuAsp 575
      1656 AATACAAATGATGCTGAAAAGATTTACTCACTTCAGTAGTAGAACATATGCTTGTCT 1715
      576 ThrAsnSerLysValIleGluGlyProGlyHisIleThrGlyValAsnLeuValTyrLeuGln 595
      1716 TCAGGTGCTTCATTTATTAAGGTCCAGATTCACAGAGGAAATTTACTATTCCTTAAA 1775
      596 -----SerGlnLysArgLeuGluIleThrCysGluThrProAsnSerThrGln 611
      1776 GATATGATTAATCATATGCTTAATTAAGTAACTTAATTAATCAGCAGCCTTGTTACAA 1835
      612 SerTyrPheIleArgLeuArgTyrAlaThrAsnGlyAlaGlyAsnThrLeuProAsnIle 631
      1836 CGATATCGGTAGTAAGATACGTATGCTTCAAC-----ACTAATCTTA 1877
      632 SerLeuThrIleProGlyValIleGlyIleProGlnArgLeuAsnAsnThrPheSer 651

```

```

Db      1878 CGACTTTTGTG-----CAAAATTCAAACATGATTTTATT 1913
      652 GlyThrAsnTyrAsn-----AsnLeuGlnTyrGlyAspPheGlyTyrPheGlnPhe 668
      1914 GTCATCTACATTAATAAATACTATGATATAGAT---GATGATTTAAACATATCAACATTT 1970
      669 ProSerThrValThrLeuProLeuAsnArgAsnIleProPhe-----682
      1971 GATTCGCACTACT-----AATCTATATATGGGTTCTCGGTCATACGAATGAA 2021
      683 -----IlePheAsnArgAlaAspValSerAsnSerIleLeuIleIleAspLysIleGlu 700
      2022 CTATATATATGAGACAGAAATCTTTCGTTCTTAATGAAAAAATCTATATATGATTAAGATAGAA 2081
      701 PheIleProIle-----704
      2082 TTTATTCCTCAGTACAAATTTGTAAGAGATTTGAAATGTAGGCGGATGCAAAATGAAAGA 2141
      705 -----ThrSerMetHisGlnAsnArgGluLys 714
      2142 ATAGGAGGTGAATTTGATGCTTAGGAAAGATTTCTTTAAGAAAAAGCAATGGAAGG 2201
      715 GlnLysLeuGluThrIleGlnThrLys---IleAsnThr 726
      2202 TATACAGTACAAATATTAAGAAATAAATTTATTAACACA 2240

RESULT 9
US-08-996-441B-37
; Sequence 37, Application US/08996441B
; Patent No. 6023013
; GENERAL INFORMATION:
; APPLICANT: English, Leigh H.
; APPLICANT: Bruseock, Susan M.
; APPLICANT: Malvar, Thomas M.
; APPLICANT: Bryson, James W.
; APPLICANT: Kulesza, Caroline A.
; APPLICANT: Walters, Frederick S.
; APPLICANT: Slatin, Stephen L.
; APPLICANT: Von Teresh, Michael A.
; APPLICANT: Romano, Charles
; TITLE OF INVENTION: INSECT-RESISTANT TRANSGENIC PLANTS
; NUMBER OF SEQUENCES: 113
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Arnold, White & Durkee
; STREET: P.O. Box 4433
; CITY: Houston
; STATE: Texas
; COUNTRY: USA
; ZIP: 77210
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/996,441B
; FILING DATE: 18-DEC-1997
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: Kitchell, Barbara S.
; REGISTRATION NUMBER: 33,928
; REFERENCE/DOCKET NUMBER: MECO:151
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 512/418-3000
; TELEFAX: 512/474-7577
; INFORMATION FOR SEQ ID NO: 37:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1959 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; FEATURE:
; NAME/KEY: CDS

```

LOCATION: 1..1956
US-08-996-441B-37

Alignment Scores:

Pred. No.:	2,586-60	Length:	1959
Score:	663.00	Matches:	208
Percent Similarity:	43.5%	Conservative:	114
Best Local Similarity:	28.1%	Mismatches:	292
Query Match:	17.1%	Indels:	126
DB:	3	Gaps:	27

US-10-783-417-2 (1-735) x US-08-996-441B-37 (1-1959)

```

QY      1 MetAenginaAsnaAspaAsnaAngluTyrgluileleAspSerHisThrSerProTyx 20
DB      1 ATGAATCCAAACAAAT---CGAAGTGAAACATGAT-----ACGATAAAGGT 42
QY      21 PheProAsnaArg-----AsnSerAsnaSerAspArgTyProTyThrAsnaProAsn 38
DB      43 ACACTTAACAGTGAATTGCAACCTAACCTAATCATATCTTTAGCTGACACATCCAAAT 102
QY      39 GluProLeuGluAsnThrAsnTyrgluTyrgluTPrleuAsnMetCysGln---GlyAsnThr 57
DB      103 TCAACACTRGAAGAATTAAATTATTAAGAAATTTTAAGATGACTGAAACAGCTTCTACG 162
QY      58 GluTyrglyAspaAsnPhenGluThrPheAlaSerAlaAspThrAlaAla---AlaValSer 76
DB      163 GAAGTGCTAGACAC-----TCTACAGTAAAGATGACAGTTGGG 201
QY      77 AlaGlyThrIleValSerGlyThrLeuLeuAlaGlyIleGlyLeuThrSerIleSer 96
DB      202 ACAGGAATTTCTGTTGTAGGCGAGATTTTGTAGGTGTAGGA----- 243
QY      97 GlyProIleGlyIleIleGlyAlaIleIleIleSerPheGlyThrLeuIleThrValPhe 116
DB      244 -----GTTCAATTTGCTGGGCACTCACTTCAATTTTATCAATCTTTCTTAACCTATA 297
QY      117 TrpProAlaGlyGluGluAsnAspLysThrValTrpThrGlnPheIleLysMetGlyGluIle 136
DB      298 TGGCCCAAGT-----GATGCTGACCGATGAAAGGCTTTTATGCGACCAAGTTGAAGTA 348
QY      137 PheValAspThrProLeuThrGluSerIleLysGlnLeuLysLeuGlnThrLeuGluGly 156
DB      349 CTGAATGATTAAGAAATAGAGAGATAGCTTAAAGTAAAGCTTCCGACAGGTATACAGGCT 408
QY      157 PheArgGlnIleLeuGlnSerTyraSerThrAlaLeuAspAspTrpArgLysLeuLysArg 176
DB      409 CTTCAAAATTAATTTGCAAGATTAATGTAATGCGTTAAATTCCTGGAAGAAA----- 459
QY      177 LeuGlnAlaProGlyLeuProProSerSerAlaLeuGlnGlnAlaAlaLeuThrLys 196
DB      460 -----ACACCTTAAAGTTTGGAGAGTAAAGAAAGCAAGCAAGATGCAATTAAGG 504
QY      197 IleArgPheGluAsnValHisAsnAspPheIleArgGluIleProGlyPheGlnLeuGlu 216
DB      505 GAACCTTTTTCACAGCAGAAAGTCAATTTCTGAATTCATCCCGCTCATTTCCAGATTTCC 564
QY      217 ThrTyrglyThrLeuLeuProIleTyrgluAlaGlnAlaAlaAsnPheHisLeuAsnLeu 236
DB      565 AAATTCGAAGTCTGTTCTACCAACATATGCAACAGGTCGCAAAATACACATTTATTTGCTA 624
QY      237 LeuGlnGlnGlyAlaGluLeuAlaAspGluTrpAsnAlaAspIleHisProSerGlnIle 256
DB      625 TTAAGAAAGTGCACAGTTTTCGAGAGAAATGGGATATTC----- 666
QY      257 GluProAsnAlaGlyThrSerAspAspTyrglyTyrglyLeuLeuLysGluAsnIleProLys 276
DB      667 -----TCGAGAGATGTTGCTGAATTTTATCATAGACATTAATAACTTACA---CAACA 717
QY      277 TyrsSerAspTyrglyAlaSerThrTyrgluThrGlyLeuLysAsnLeuArgAspGluPro 296
DB      718 TACACTGACCATGTTGTTAATTGTTAAATGTTGATTAATATGTTTAAAGAGTTCAACT 777

```

```

QY      297 AsnMetLysTrpSerIlePheAsnaAspTyrgluArgTyrgluMetThrIleThrValLeuAsp 316
DB      778 TATGATGATGGGTCAATTTTAAACGTTTTCGACAGAAATGACTTAATCTATTAATGAT 837
QY      317 ThrIleSerGlnPheSerLeuTyraAspIleLysArgTyrgluArgAspSerIleGlyGlyIle 336
DB      838 CTAATTTGACTTTTCCCATTTTATGATATATGCGTTATCTCA----- 879
QY      337 GluValLysGlyIleLysAsnGluLeuThrArgGluIleTyrgluThrGluIleAsnPhe 356
DB      880 -----AAAGCGTTAAACAGAACTAACAGAGCATTTTACGATTCATATTTTCT 933
QY      357 AspArgLeuProGlnLeuArgValGlnProAsnLeuAlaThrMetGluTyraAsnLeuThr 376
DB      934 CTTAGGACACCA---CTTGCCTACCGACCAACTTTTGTAGATATGAATAACTATTCGA 990
QY      377 ArgAlaSerPheLysLeuPheSerPheLeuGluGlnPheIlePheTyThr 393
DB      991 AAACCT-----CATTTATTTGATTATTATACAGGGATTTGAATTTCAATACGCTTCA 1044
QY      394 -----GluAsnThrAsnPhe-----GlyAsnArgLeuValGlyIle 405
DB      1045 CTGTGTTACTTTGGGAAAGATTTCTTCAATTAATGCTGTGTAATTA-----GTA 1095
QY      406 SerAsnaArgAspAlaProThrTyrsSerAsnThrIleThrGluThrLeuTyrglyGluArg 425
DB      1096 GAACCTAGACCTAGATATGAGATCTATGAGACAAATTAATCTCCCATTTTATGAGATATA 1155
QY      426 ThrGlySerProThrThrIleTyThrIleArgProPheGluSer-----TyrglyVal 442
DB      1156 TCTACTGAACCTGTACAAACCTA-----AGCTTGTATGACAAAGTTTATCGAACT 1209
QY      443 SerIleValThrAspArgGlnSerProPro-----ValSerPro 455
DB      1210 ATAGCTATATACAGAGCTAGCGGCTTGGCCGAATGTAAGTATATTAGCTTACGAAA 1269
QY      456 IleGlnProHisPheIleIleAsnGlnIleGluLeuTyrglyLeuAsnGlySerSerAsn 475
DB      1270 GTTGAT-----TTTAGTCAATATGATGATCAAAAAGAACTAGTACACAA 1317
QY      476 ThrLeuLysTyrsSerAlaGlyLysLeuSerAsnTyrgluAsnThrThrPheGln 495
DB      1318 ACATATGATTCAAAAGAAACATGCGCATGTAATGACAGGATTTATGACCAATTA 1377
QY      496 PheProArgLysAspCysAsnLeuValIleAspProGlyCysSerProAsnPheAsn 515
DB      1378 CCGCCAGAAACACAGAT-----GAACCACTGAAAAA 1410
QY      516 AsnTyrsSerHisIleLeuSerHisPheSerLeuPheThrTyrsTyrgluIleGlyLeu 535
DB      1411 GCATATAGCTATCAGCTTAATTAACCGGAAATGTTCTTA-----ATG 1452
QY      536 GlnLeuGlnIleLeuAspThrGlyValLeuGlyTyrgluThrHisSerSerValAspArgTyrg 555
DB      1453 CAGAGCCGTCGTGAAACATTCATTTTAACTTGTGACACATGAAAGGTATACATTTT 1512
QY      556 AsnAlaIleSerAspLysIleIleThrMetCileProAlaIleLysGlyAsnaAsnLeuAsp 575
DB      1513 AATACAAATTAATGCTGAAAAGATTTACTCACTTCAGATGAGAAAGCATATGCCCTTGCT 1572
QY      576 ThrAsnSerLysValIleGluGlyProGlyHisThrGlyGlyAsnLeuValTyrgluGln 595
DB      1573 TCAGGTGCTTCATATTAAGAGTCCAGAGATTCACAGAGGAAATTTACTATCTTA 1632
QY      596 -----SerGlnGlyArgLeuGluIleThrCysGluThrProAsnSerThrGln 611
DB      1633 GAATCTAGTAATTCATATGCTTAATTTAAAGTTATCAATTAATTCAGAGCCCTTGTTACA 1692
QY      612 SerTyrgluIleArgLeuArgTyrgluAlaSerThrTyrgluAlaSerThrLeuProAsn 631
DB      1693 CGATATCGTGTAGAAATAGGCTATCTCTAC-----ACTAATCTTA 1734
QY      632 SerLeuThrIleProGlyValIleGlyIleProProGlnArgLeuAsnaAsnThrPheSer 651

```

```

Db      1735 CGACCTTTTGTG-----CAAAATTCAACATGATTTTCTT 1770
QY      652 GYThrAsnTyrAsn-----AsnLeuGlnTyrGlyAspPheGly 664
Db      1771 GTGATCTACATTAAATAAATGATGAATGAATGATGAATCAATTAATGAAT 1830
QY      665 TyrPheGlnPheProSerThrValThrLeuProLeuAsnArgAsnIleProPheIlePhe 684
Db      1831 CTCGCACTACTAATATTCTAATATGGGTTCTCGGGGATGATGATGATGATGATGATGAT 1890
QY      685 AsnArgAlaAspValSerAsnSerIleLeuIleIleAspLysIleGluPheIleProIle 704
Db      1891 GCAGAACTCTTGCTTTCTAATGAAAAAATCTATATGATGATGATGATGATGATGATGAT 1950

RESULT 10
US-08-993-722A-37
/ Sequence 37, Application US/08993722A
/ Patent No. 6060594
/ GENERAL INFORMATION:
/ APPLICANT: English, Leigh H.
/ APPLICANT: Brubock, Susan M.
/ APPLICANT: Malvar, Thomas M.
/ APPLICANT: Bryson, James W.
/ APPLICANT: Kulcsza, Caroline A.
/ APPLICANT: Walters, Frederick S.
/ APPLICANT: Staltn, Stephen L.
/ APPLICANT: Von Tersech, Michael A.
/ TITLE OF INVENTION: NUCLEIC ACID SEGMENTS ENCODING MODIFIED
/ TITLE OF INVENTION: COLEOPTERAN-TOXIC CRYSTAL PROTEINS
/ NUMBER OF SEQUENCES: 113
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Arnold, White & Durkee
/ STREET: P.O. Box 4433
/ CITY: Houston
/ STATE: Texas
/ COUNTRY: USA
/ ZIP: 77210
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: Patent Release #1.0, Version #1.30
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/993,722A
/ FILING DATE: 18-DEC-1997
/ CLASSIFICATION: 435
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Kitchell, Barbara S.
/ REGISTRATION NUMBER: 33,928
/ REFERENCE/DOCKET NUMBER: MECO:149
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 512/418-3106
/ TELEFAX: 512/474-7577
/ INFORMATION FOR SEQ ID NO: 37:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 1959 base pairs
/ TYPE: nucleic acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ FEATURE:
/ NAME/KEY: CDS
/ LOCATION: 1..1956
/ US-08-993-722A-37

Alignment Scores:
Pred. No.: 2 586-60 Length: 1959
Score: 653.00 Matches: 208
Percent Similarity: 43.5% Conservative: 114
Best Local Similarity: 28.1% Mismatches: 292
Query Match: 17.1% Indels: 126
DB: 3 Gaps: 27

```

```

US-10-783-417-2 (1-735) x US-08-993-722A-37 (1-1959)
QY      1 MetAsnGlnAsnAsnAsnAsnGlnTyrGluIleIleAspSerHisThrSerProTyr 20
Db      1 ATGAATCCAAACAAAT---CGAAGTAAACATGAT-----ACGATAAAGATT 42
QY      21 PheProAsnArg-----AsnSerAsnAspSerArgTyrProTyrThrAsnAsnProAsn 38
Db      43 ACACCTTACAGGAATTTGCAAACTTAACCTATATCAATATCTTTAGCTGACAAATCCAAAT 102
QY      39 GlnProLeuGlnAsnThrAsnTyrLysGlnTyrPheLeuAsnMetCysGln---GlyAsnThr 57
Db      103 TCACACCTAGCAAGAAATTAATTAATAAGAAATTTTAAGAAATGACGCAAGACAGCTTCTACG 162
QY      58 GlnTyrGlyAspAsnPheGluThrPheAlaSerAlaSerIleAla---AlaValSer 76
Db      163 GAAGTGTAGACAAAC-----CTACAGTAAAGATGACAGTGTGG 201
QY      77 AlaGlyThrIleValSerGlyThrLeuLeuAlaGlyIleGlyGlyLeuThrSerIleSer 96
Db      202 ACAGGAATTTCTGTTGTAGGCGACATTTTAGGTGTAGGA----- 243
QY      97 GlyProIleGlyIleIleGlyAlaIleIleIleSerPheGlyThrLeuIleThrValPhe 116
Db      244 -----GTTCCATTTGCTGGGCGACCTCACTTATTTATCATCTTTCTTAACACTATA 297
QY      117 TrpProAlaGlyGlnGlnAspLysThrValThrPheGlnPheIleLysMetGlyGluIle 136
Db      298 TGGCCCAAGT-----GATCTGACCCATGAAAGCGCTTTATGTCACAAAGTTGCAAGTA 348
QY      137 PheValAspThrProLeuThrGluSerIleLysGlnLeuLysLeuGlnThrLeuGlnGly 156
Db      349 CTGATGATAGAAATAGAGAGATGATGCTAAAGTAAAGCTTTCGACAGATTACGGGT 408
QY      157 PheArgGlnIleLeuGlnSerTyrAsnThrAlaLeuAspAspTyrArgLysLeuLysArg 176
Db      409 CTTCAAATTAATTTGGAAGATTATGTTAATGCTTAAATTCCTGGAAGAA----- 459
QY      177 LeuGlnAlaProGlyLeuProProSerSerAlaLeuGlnAlaAlaLeuThrLeuLys 196
Db      460 -----ACACCTTAAGTTGCGAAGTAAAGAAAGCAAGATGCAATGAATAGG 504
QY      197 IleArgPheGlnAsnValHisAsnAspPheIleArgGluIleProGlyPheGlnLeuGlu 216
Db      505 GAACCTTTTCTCAAGCAGAAAGATGCTAATTCGTAATTCGATCCGCTCATTTCCAGTTTCC 564
QY      217 ThrTyrLysThrLeuLeuLeuProIleTyrAlaGlnAlaAlaAsnPheHisLeuAsnLeu 236
Db      565 AAATTCGAAGTCTGTTTCTACCAACATATGCAAGCTGCAAAATACATTTATTTGCTA 624
QY      237 LeuGlnGlnGlyAlaGluLeuAlaAspGluTyrPheAsnAlaAspIleHisProSerGlnIle 256
Db      625 TTTAAAGATGCTCAAGTTTTCGACAGAAATGGGATATTC----- 666
QY      257 GluProAsnAlaGlyThrSerAspAspTyrTyrLysLeuLeuLysGluAsnIleProLys 276
Db      667 -----TCAGAAAGATGTCGAATTTATATCAATACAAATTAACCTTAACA---CAACAA 717
QY      277 TyrSerAsnTyrCysAlaAsnThrTyrArgThrGlyLeuLysAsnLeuArgAspGluPro 296
Db      718 TACACTGACCATGTTGTTAATTTGATTAATGTTGAATTAATGCTTTAAGAGTTCAACT 777
QY      297 AsnMetLysTrpSerIlePheAsnAspTyrArgArgTyrMetThrIleThrValLeuAsp 316
Db      778 TATGATGATGCGTCAAAATTTAACCCTTTTCGACAGAAATGACCTTAACCTTATTAAGAT 837
QY      317 ThrIleSerGlnPheSerLeuTyrAspIleLysArgTyrArgAspSerIleGlyGlyIle 336
Db      838 CTAAATTTGACTTTTCCATTTATATATATGCTGATTACTCA----- 879
QY      337 GluValLysGlyIleLysAsnGluLeuThrArgGluIleTyrThrThrGluIleAsnPhe 356

```


OY	39	GLNProLeuGlnAAsnThrAsnTYrLeuGluTYrProLeuAsnMetCysGln----	GLYAsnThr	57
Db	103	TCACACCTAGAGAAATTAATTAATAAGAAATTTTAAGAAGTAACTGAGGACAGCTTACG	----	162
OY	58	GlnTYrGlyAspAsnPhenGluThrPheAlaSerAlaAspThrIleAla----	AlaValSer	76
Db	163	GAGTGCCTAGACAAAC-----TCACAGTAAAGAGTCACTTGG	----	201
OY	77	AlaGlyThrIleValSerGlyThrIleuAlaGlyIleGlyGlyLeuThrSerIleSer	96	
Db	202	ACAGGAATTTCTGTTGTAGAGGACAGATTATAGCTGTGTAGGA-----	----	243
OY	97	GlyProIleGlyIleIleGlyAlaIleIleIleSerPheGlyThrIleuIleThrValPhe	116	
Db	244	-----GTTCAATTTCTGGGGCAGCTCACTCATTTTATCAATCTTTCTTAACCTATA	297	
OY	117	TrpProAlaGlyGluGlnAspLysThrValTrpThrGlnPheIleLysMetGlyIle	136	
Db	288	TGGCCAACT-----GATGCTAACCCATGGAAGGCTTTATGGCACAAGTTGAATA	348	
OY	137	PheValAspThrProLeuThrGluSerIleLysGlnLeuLysLeuGlnThrIleuGlu	156	
Db	349	CTGATAGATTAAGAAATATGAGAGATGCTAAATGAATGAAGCTTCGACAGATTACAGGT	408	
OY	157	PheArgGlnIleLeuGlnSerTYrAsnThrAlaLeuAspAspTrpArgLysLeuLysArg	176	
Db	409	CTTCAAAATTAATTTGGAAGATTATTATAGGCTTAAATTCCTGGAATA-----	459	
OY	177	LeuGlnAlaProGlyLeuProProSerSerAlaLeuGlnAlaIleThrIleuLys	196	
Db	460	-----ACACCTTAAGTTTGGCAAGTAAAGAGCCAAAGATGAAATAGG	504	
OY	197	IleArgPheGluAsnValHisAsnAspPheIleArgGluIleProGlyPheGlnLeu	216	
Db	505	GACCTTTTTCACAGCAAGAAAGTCATTTTGCTATTCATGCGCTCATTTGCAGTTCC	564	
OY	217	ThrTYrLysThrLeuLeuLeuProIleTYrAlaGlnAlaIleAsnPheHisLeuAsnLeu	236	
Db	565	AAATTGCAAGCTGCTTTCTTCCACCAATATGCAACAGCTGCMAATTAACATTATTTGCTA	624	
OY	237	LeuGlnGlnGlyAlaGluLeuAlaAspGluTrpAsnAlaAspIleHisProSerGlnIle	256	
Db	625	TTAAAGATGCTCAAGCTTTTGGAGAAAGATGGGATTTCT-----	666	
OY	257	GluProAsnAlaGlyThrSerAspAspTYrTYrLysLeuLeuLysGluAsnIleProLys	276	
Db	667	-----TCAGAAAGATGTTGCTGAATTTATATCAATACATTTAAACTTACA---CAACA	717	
OY	277	TYrSerAsnTYrCysAlaAsnThrTYrArgThrGlyLeuLysAsnLeuArgAspGluPro	296	
Db	718	TACACTGACCAATGTGTTAATGTGATATAATGTGATTAATAATGTTTAAAGAGTTCACT	777	
OY	297	AsnMetLysTrpSerIlePheAsnAspTYrArgArgTYrMetThrIleThrValLeuAsp	316	
Db	778	TATGATGATGGGTCAATTTAAACCGTTTTCGACAGAAATGACTTTAACTGATTAATAT	837	
OY	317	ThrIleSerGlnPheSerLeuTYrAspIleLysArgTYrArgAspSerIleGlyIle	336	
Db	838	CTAATTTGACTTTTCCATTTATATATATTCGTTATATCTA-----	879	
OY	337	GluValLysGlyIleLysAsnGluLeuThrArgGluIleTYrThrThrGluIleAsnPhe	356	
Db	880	-----AAAGGGGTTAAACAGAACTTAAACAGACAGCATTTTAAACGATTCATTTTCT	933	
OY	357	AspArgLeuProGlnLeuArgValGlnProAsnLeuAlaThrMetGluTYrAsnLeuThr	376	
Db	934	CTTAGACACCA---CTTGCGTAGCGACCACTTTTGGAGTATAGAAACTTATTCGA	990	
OY	377	ArgAlaSerPheLysLeuPheSerPheLeuGlnGlnPheIlePheTYrThr-----	393	
Db	991	AAACT-----CATTATTTGATATTATTAACGGGAGTTGAATTCATACGGCTTCA	1044	
OY	394	-----GluAsnThrAsnPhe-----GlyAsnArgLeuValGlyIle	405	

Db 1045 CCTGGTACTTGGGAAAGATCTTTCATTTATGCTGCTGTAATTAT-----GTA 1095
QY 406 SerAsnArgAspAlaProThrTySerAsnThrIleThrGluThrLeuTyArgIleuArg 425
Db 1096 GAAACTGACCTAGCTATAGAGATCTACAGTAGACATTAATCTCCCATTTTATGGAGATAA 1155
QY 426 ThrGlySerProThrThrIleuTyThrIleArgProGluGluSer-----TyLeuVal 442
Db 1156 TCTACTGACCTGTACAAACACTA-----AGCTTTATGACAAAAAAGTTTATGCAACT 1209
QY 443 SerIleValThrAspArgGlnSerProPro-----ValSerPro 455
Db 1210 ATAGCTATATACAGCGTAGAGCGGCTGGCCGAAGGTAAAGGATATATTAGGTTTACGAA 1269
QY 456 IleGlnProHisPheIleIleAsnGlnIleGluLeuTyIleuAsnGlySerSerAsnAsn 475
Db 1270 GTTGAAT-----TTTAGTCATATAGATATATCAAAAAATGCAACTGATGACAA 1317
QY 476 ThrLeuTyTySerAlaGlyIleSerLeuSerAsnTyIleGlnThrThrPhePheGln 495
Db 1318 ACATATATGATTTCAAAAAAAGAAACATGCGCATGTATAGTCACAGATTTCTATTGACCAATTA 1377
QY 496 PheProArgIleuTyIleuAspCysAsnLeuValIleAspProGlyCysSerProAsnPheAsn 515
Db 1378 CCGCCAGAAACAAACAGAT-----GAACCACTTGAAAAA 1410
QY 516 AsnTyTySerHisIleLeuSerHisPheSerLeuPheThrTySerTyIleValIleGlyLeu 535
Db 1411 GCATATATGATCTACAGCTTAATTTAGCGCAATGTTTCTTA-----ATG 1452
QY 536 GlnLeuGlnIleLeuAspThrGlyValIleuGlyTyTrpThrHisSerSerValAspArgTy 555
Db 1453 CAGACCGCGTGGAAACAATTCATTTTATCTTGACACATTAAGATGATGACATTTT 1512
QY 556 AsnAlaIleSerAspIleIleIleThrMetIleProAlaIleIleGlyAsnAsnLeuAsp 575
Db 1513 AATCAATTTGATGCTGAAAGATTAATTCATCTCACTAGTAGAAAGCATATGCGCTTCT 1572
QY 576 ThrAsnSerTyValIleGluGlyProGlyValHisPheGlyValAsnLeuValTyIleuGln 595
Db 1573 TCAGGTGCTTCATTTATTTAGAGGTCGAGATTCACAGAGAAATTTTACATTTCTTAAAA 1632
QY 596 -----SerGlnGlyArgLeuGluIleThrCysGluThrProAsnSerThrGln 611
Db 1633 GAATCTAGTATTCATCAATTCGTAATTTTAAAGTTCATTAATTAATTCAGACGCTTGTACAA 1692
QY 612 SerTyTyPheIleArgLeuArgTyValIleThrAsnGlyValGlyAsnThrIleuProAsnIle 631
Db 1693 CGAATATCGTGAAGAAATGCGTATAGCTTCAAC-----ACTAATCTTA 1734
QY 632 SerLeuThrIleProGlyValIleGlyIleProProGlnArgLeuAsnAsnThrPheSer 651
Db 1735 CGACTTTTGTG-----CAAATTCAAACATGATTTCTT 1770
QY 652 GlyThrAsnTyAsn-----AsnLeuGlnTyArgIleAspPheGly 664
Db 1771 GTCATCTACATTAATAAAACCTAGTAATTAAGATGATTAATTAATTCATTCACAAACCTTGAT 1830
QY 665 TyTyPheGlnPheProSerThrValThrIleuProLeuAsnArgAsnIleProPheIlePhe 684
Db 1831 CTCGCACTACTAATTTCTAATATGGGGTCTCCGGGTGATATGAATGAACCTTATATATGGA 1890
QY 685 AsnArgAlaAspValSerAsnSerIleuLeuIleAspTyIleGluPheIleProIle 704
Db 1891 GCAGATATCTTTCGTTTCAATGAAAAAATATCATATATAGATATGATGATATATATCCACGTA 1950

RESULT 12
US-08-993-775B-37
; Sequence 3, Application US/08993775B
; Patent No. 6077824
; GENERAL INFORMATION:
; APPLICANT: English, Leigh H.


```

Qy 443 SerIleValThrAspArgInserPro-----ValSerPro 455
Db 1210 ATAGCAATACACAGCTAGCGGCTGCGCAATGTAAGTATATTAGGTGTACGAA 1269
Qy 456 IleGlnProHisPheIleIleLeuGlnIleGluLeuTyrLeuAsnGlySerSerAsn 475
Db 1270 GTTGAAT-----TTTAGTCAATATGATGATCAAAAATGAACTGATACAA 1317
Qy 476 ThrLeuLysTyrSerAlaGlySerLeuSerAsnTyrGlnAsnThrPhePheGln 495
Db 1318 ACATATGATTCAGAAAGAAACAAATGCGCATGTAAGTCAAGATTCATGTCACAA 1377
Qy 496 PheProArgLysLeuAspCysAsnLeuValIleAspProGlyCysSerProAsn 515
Db 1378 CCGCCAGAAACAAACAT-----GACCACTTGAAAA 1410
Qy 516 AsnTyrSerHisIleLeuSerHisPheSerLeuPheThrTyrSerTyrValIleGlu 535
Db 1411 GCATATAGTCATCAGCTTATATACGCGAATGTTCTTA-----ATG 1452
Qy 536 GlnLeuGlnIleLeuAspThrGlyValLeuGlyTyrThrHisSerSerValAspArg 555
Db 1453 CAGGACCGTCGTGGAACAAATTCATTTTACTTGACACATAGAGTGTAGCTTTT 1512
Qy 556 AsnAlaIleSerAspLysIleIleThrMetIleProAlaIleLysGlyAsnAsnLeu 575
Db 1513 AATACATTCATGCTGAAAAAGATTACTCAATTCCTCGATGTAACATATGCTTGT 1572
Qy 576 ThrAsnSerLysValIleGluGlyProGlyHisThrGlyValAsnLeuValTyrLeu 595
Db 1573 TCAGGTGCTTCATATTCAGAGTCCAGATTCACAGAGAAATTCATATCTCAAAA 1632
Qy 596 -----SerGlnLysArgLeuGluIleThrCysGluThrProAsnSerThrGln 611
Db 1633 GAATCTAGTAAATTCATTCCTAAATTTAAAGTTACCTTAATTCACAGCTTTGCA 1692
Qy 612 SerTyrPheIleArgLeuArgTyrAlaThrAsnGlyValGlyAsnThrLeuProAsn 631
Db 1693 CCATATCGGTATGAAATACGCTATGCTTCTACC-----ACTAATCTTA 1734
Qy 632 SerLeuThrIleProGlyValIleGlyIleProProGlnArgLeuAsnAsnThrPhe 651
Db 1735 CCACCTTTTGTG-----CAAAATTCAAACATGATTTTCT 1770
Qy 652 GlyThrAsnTyrAsn-----AsnLeuGlnTyrGlyAspPheGly 664
Db 1771 GTCATCTACATTAATTAACATATGAATGAATGAATGATGATTAACATATTCAT 1830
Qy 665 TyrPheGlnPheProSerThrValThrLeuProLeuAsnArgAsnIleProPheIle 684
Db 1831 CTCGCACTACTTAATTCATATGAGGCTTCTCGGATGATGAATGAATCTTATTA 1890
Qy 685 AsnArgAlaAspValSerAsnSerIleLeuIleIleAspLysIleGluPheIlePro 704
Db 1891 GCAGATCTTGTGTTCTAATGAAAAAATCTAATATGATGATGATTAATTCAGTA 1950

```

RESULT 13
US-09-427-770-37
Sequence 37, Application US/09427770
Patent No. 6620988

GENERAL INFORMATION:

APPLICANT: English, Leigh H.
APPLICANT: Brusock, Susan M.
APPLICANT: Malvar, Thomas M.
APPLICANT: Bryson, James W.
APPLICANT: Kulesza, Caroline A.
APPLICANT: Walters, Frederick S.
APPLICANT: Slatin, Stephen L.
APPLICANT: Von Terssch, Michael A.
TITLE OF INVENTION: NUCLEIC ACID SEGMENTS ENCODING MODIFIED
TITLE OF INVENTION: COLEOPTERAN-TOXIC CRYSTAL PROTEINS
NUMBER OF SEQUENCES: 113

```

CORRESPONDENCE ADDRESS:
ADDRESSER: Arnold, White & Durkee
STREET: P.O. Box 4433
CITY: Houston
STATE: Texas
COUNTRY: USA
ZIP: 77210
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/427,770
FILING DATE:
CLASSIFICATION:
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US 08/993,722
FILING DATE: 18-DEC-1997
ATTORNEY/AGENT INFORMATION:
NAME: Kitchell, Barbara S.
REGISTRATION NUMBER: 33,928
REFERENCE/DOCKET NUMBER: MECO:149
TELEPHONE: 512/418-3106
TELEFAX: 512/474-7577
INFORMATION FOR SEQ ID NO: 37:
SEQUENCE CHARACTERISTICS:
LENGTH: 1959 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
FEATURE:
NAME/KEY: CDS
LOCATION: 1..1956
US-09-427-770-37
Alignment Scores:
Pred. No.: 2,586-60 Length: 1959
Score: 663.00 Matches: 208
Percent Similarity: 43.5% Conservative: 114
Best Local Similarity: 28.1% Mismatches: 292
Query Match: 17.1% Indels: 126
Gaps: 27
US-10-783-417-2 (1-735) x US-09-427-770-37 (1-1959)
Qy 1 MetAsnGlnAsnAspAsnAsnGluTyrGluIleIleAspSerHisThrSerProTyr 20
Db 1 ATGAATCCAAACAT---CGAAGTGAACATGAT-----ACGATTAAGTT 42
Qy 21 PheProAsnArg-----AsnSerAsnAspSerArgTyrProTyrThrAsnProAsn 38
Db 43 ACACCTTAACAGTGAATGCAATCAATCAATATCAATATCTTAGTGAATCCAAAT 102
Qy 39 GlnProLeuGlnAsnThrAsnTyrIleGluTyrIlePheLeuAsnMetCysGln---GlyAsnThr 57
Db 103 TCACACCTAGAGAAATTAATTAAGAAATTTTAAGATTAACGACAGGTTCTAGC 162
Qy 58 GlnTyrGlyAspAsnPheGluThrPheAlaSerAlaAspThrIleAla---AlaValSer 76
Db 163 GAAGTCTAGACAC-----TCAACGTAAGAAATGCAAGTTCGG 201
Qy 77 AlaGlyThrIleValSerGlyThrLeuLeuAlaGlyIleGlyGlyLeuThrSerIleSer 96
Db 202 ACAGGAATTCCTGTGTAGGCGAGATTTTAGGTGTGGA----- 243
Qy 97 GlyProIleGlyIleIleGlyAlaIleIleIleSerPheGlyThrLeuIleThrValPhe 116
Db 244 -----GTTCAATTCGTGGGCGACTCACTTCATTTTATCAATCAATTCCTTAAC 297
Qy 117 TTPProAlaGlyGluGlnAspLysThrValTTPThrGlnPheIleLysMetGlyGluIle 136

```

Db 238 TGGCCAGT-----GATGCTGACCCAGTGAAGGCTTTATATG3CACAACTGAAGTA 348
 Qy 137 PheValAspThrProLeuThrGluSerIleValGlnLeuValLeuGlnThrLeuGluGly 156
 Db 349 CTGATAGATTAAGAAATAGAGATATGCTAAAGTAAAGCTCTTGCGACAGTTACAGGCT 408
 Qy 157 PheArgGlnIleLeuGlnSerTyrAsnThrAlaLeuAspAspTyrArgIleValSerArg 176
 Db 409 CTTCAAAATTAATTTGCAAGATTAATGATGCTTAAATTCCTGGAGAGAAA----- 459
 Qy 177 LeuGlnAlaProGlyLeuProProSerSerAlaLeuGlnGlnAlaAlaLeuThrLeuVal 196
 Db 460 -----ACACCTTTAAGTTTGCAGAGTAAAGAGCCAGACATGCAATTAAGG 504
 Qy 197 IleArgPheGluAsnValHisAsnAspPheIleArgGlnIleProGlyPheGlnLeuGlu 216
 Db 505 GAACCTTTTCTCAACGACAGAAAGTCTTTTGTATATTCATGCGCATTTGCACTTCC 564
 Qy 217 ThrTyrLeuThrLeuLeuLeuProIleTyrAlaGlnAlaAlaAsnPheHisLeuAsnLeu 236
 Db 565 AAATTCGAAGTCTGTTCTTACCAACATATGCAAGCTGCAAAATACACATTTATGCTA 624
 Qy 237 LeuGlnGlnGlyAlaGluLeuAlaAspGluTyrAsnAlaAspIleHisProSerGlnIle 256
 Db 625 TTAAAGATGCTCAAGTTTTCGAGAGAGATGCGGATATCT----- 666
 Qy 257 GluProAsnAlaGlyThrSerAspAspTyrTyrIleLeuLeuValGlnAsnIleProVal 276
 Db 667 -----TCGAAGATGTTGCTGAATTTATCATAGACATTAAGAACTTACA---CAACA 717
 Qy 277 TyrSerAsnTyrCysAlaAsnThrTyrArgThrGlyLeuValAsnLeuArgAspGluPro 296
 Db 718 TACACTGCACATGTTGTATATGTAATGTTGATTAATGATGTTAAAGAGTTCAACT 777
 Qy 297 AsnMetLysTyrSerIlePheAsnAspTyrArgArgTyrMetThrIleThrValLeuAsp 316
 Db 778 TATGATGATGGGTCAAAATTTAACCGTTTTCGACAGAAATGACTTAACTGTATTAAT 837
 Qy 317 ThrIleSerGlnPheSerLeuTyrAspIleLysArgTyrArgAspSerIleGlyGlyIle 336
 Db 838 CTAAATTTGATCTTTCCCATTTATATGATATTCGGTATACTCA----- 879
 Qy 337 GluValIleGlyIleLysAsnGlnLeuThrArgGlnIleTyrThrThrGlnIleAsnPhe 356
 Db 880 -----AAAGGGGTAAACAGAACATCAACAGACATTTTACGATCCAAATTTTTCT 933
 Qy 357 AspArgLeuProGlnLeuArgValGlnProAsnLeuAlaThrMetGluTyrAsnLeuThr 376
 Db 934 CTTAGGACACCA---CTTGCCTACGACCACTTTTGTGATATGAAAACTCTATTCGA 990
 Qy 377 ArgAlaSerPheLysLeuPheSerPheLeuGlnGlnPheIlePheTyrThr----- 393
 Db 991 AAACCT-----CATTTATTTGATATTTTACAGGGATTTGAATTTCAATCGGCTTCA 1044
 Qy 394 -----GluAsnThrAsnPhe-----GlyAsnArgLeuValGlyIle 405
 Db 1045 CTGTGACTTTGGAGAAAGATTTCTTCAATATATGCTGTGTAATAT-----GTA 1095
 Qy 406 SerAsnArgAspAlaProThrTyrSerAsnThrIleThrGlnThrLeuTyrGlyGluArg 425
 Db 1096 GAACCTAGACCTAGATAGATCTAGTAAAGACATTAATCTCCCATTTTATAGAGATTA 1155
 Qy 426 ThrGlySerProThrThrIleThrIleArgProPheGluSer-----TyrLysVal 442
 Db 1156 TCTACTGAACCTGTACAAAAGCTA-----ACCTTGTATGACAAAAAGTTATGCACT 1209
 Qy 443 SerIleValThrAspArgLysProPro-----ValSerPro 455
 Db 1210 ATAGCTAATATACAGAGTACGCTGGCGGAATGGAAGTATATTAGGTATACGAAA 1269
 Qy 456 IleGlnProHisPheIleIleAsnGlnIleGluLeuTyrLeuAsnGlySerSerAsnAsn 475
 Db 1270 GTTGTAT-----TTTAGTCATATGATGATCAAAAAAATGAATGAACTAGTACACA 1317

Qy 476 ThrLeuLysTyrSerAlaGlyGlySerLeuSerAsnTyrGlnAsnThrThrPhePheGln 495
 Db 1318 ACATATGATTTCAAAAAAGAAACAAATGCCATGTAAGTACACAGGATTTATGACCAATTA 1377
 Qy 496 PheProArgLysValAspCysAsnLeuValIleAspProGlyCysSerProAsnPheAsn 515
 Db 1378 CCGCCAGAAACAAACAGAT-----GAACCACTTGAAAAA 1410
 Qy 516 AsnTyrSerHisIleLeuSerHisPheSerLeuPheThrTyrSerTyrValIleGlyLeu 535
 Db 1411 GCATATATGATCATCAGCTTAATTAATTCGCGGAATGTTCTTA-----ATG 1452
 Qy 536 GlnLeuGlnIleLeuAspThrGlyValLeuGlyIleTyrThrHisSerSerValAspArgTyr 555
 Db 1453 CAGAGCCGTGCGGAAACATTCATTTTACTTGACACATAGAAAGTATGACTTTT 1512
 Qy 556 AsnAlaIleSerAspLysIleIleThrMetIleProAlaIleLysGlyAsnAsnLeuAsp 575
 Db 1513 AATACAAATTTGATGCTGAAAGATTAATCACTTCCAGTATGTAAGCATATGCCCTGTCT 1572
 Qy 576 ThrAsnSerLysValIleGluGlyProGlyHisThrGlyValAsnLeuValTyrLeuGln 595
 Db 1573 TCAGTGTCTTCATTTATGAAAGTCCAGGATTCACAGAGAAATTTACTATTCCTTAAA 1632
 Qy 596 -----SerGlnGlyArgLeuGlnIleThrCysGluThrProAsnSerThrGln 611
 Db 1633 GAATCTAGTAATTCATTAATTAAGTAAATTAAGTAAATTAATTCAGACGCTTGTTACA 1692
 Qy 612 SerTyrPheIleArgLeuArgTyrAlaThrAsnGlyAlaGlyAsnThrLeuProAsnIle 631
 Db 1693 CGATATCGTGAAGATATGCTATGCTCTTCC-----ACTAACTTA 1734
 Qy 632 SerLeuThrIleProGlyValIleGlyIleProProGlnArgLeuAsnAsnThrPheSer 651
 Db 1735 GCACCTTTTGTG-----CAAAATTCAAACATGATTTTCTT 1770
 Qy 652 GlyThrAsnTyrAsn-----AsnLeuGlnTyrGlyAspPheGly 664
 Db 1771 GTCATCTAATTAATTAAGATATGATATGAATGATGATTTTAACATTAACAAATTTGAT 1830
 Qy 665 TyrPheGlnPheProSerThrValThrLeuProLeuAsnAspAsnIleProPheIlePhe 684
 Db 1831 CTCGCACACTAATTAATTAATGAGGTCTCGGGTGAATGAAGAACTTATTAATGGA 1890
 Qy 685 AsnArgAlaAspValSerAsnSerIleLeuIleIleAspLysIleGlyPheIleProIle 704
 Db 1891 GCAGATCTCTTCGTTCTTAATGAAAAAATCTATATAGATAGATTAATTTATCCAGTA 1950
 RESULT 14
 US-09-427-769-37
 ; Sequence 37, Application US/09427769
 ; Patent No. 6642030
 ; GENERAL INFORMATION:
 ; APPLICANT: English, Leigh H.
 ; APPLICANT: Bruesock, Susan M.
 ; APPLICANT: Malvar, Thomas M.
 ; APPLICANT: Bryson, James W.
 ; APPLICANT: Kuleza, Caroline A.
 ; APPLICANT: Walters, Frederick S.
 ; APPLICANT: Slatin, Stephen L.
 ; APPLICANT: Von Teresh, Michael A.
 ; APPLICANT: Romano, Charles
 ; TITLE OF INVENTION: NUCLEIC ACID SEGMENTS ENCODING MODIFIED
 ; TITLE OF INVENTION: COLEOPTERAN-TOXIC CRYSTAL PROTEINS
 ; NUMBER OF SEQUENCES: 113
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSER: Arnold, White & Durkee
 ; STREET: P. O. Box 4433
 ; CITY: Houston
 ; STATE: Texas
 ; COUNTRY: USA
 ; ZIP: 77210

```

COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/427,769
FILING DATE:
CLASSIFICATION:
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: 08/993,722
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Kitchell, Barbara S.
REGISTRATION NUMBER: 33,928
REFERENCE/DOCKET NUMBER: MECO:149
TELECOMMUNICATION INFORMATION:
TELEPHONE: 512/418-3106
TELEFAX: 512/474-7577
INFORMATION FOR SEQ ID NO: 37:
SEQUENCE CHARACTERISTICS:
LENGTH: 1959 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
FEATURE:
NAME/KEY: CDS
LOCATION: 1..1956
US-09-427-769-37
Alignment Scores:
Pred. No.: 2.58e-60 Length: 1959
Score: 663.00 Matches: 208
Percent Similarity: 43.5% Conservative: 114
Best Local Similarity: 28.1% Mismatches: 292
Query Match: 17.1% Indels: 126
Gaps: 27
DB: 3
US-10-783-417-2 (1-735) x US-09-427-769-37 (1-1956)
QY 1 MetcAAGlnAsnAspAsnAsnGluTyrGluIleIleAspSerHisThrSerProTyr 20
DB 1 AAGATTCACAAACAT---CGAAGTGAACATGAT-----ACGATTAACGTT 42
QY 21 PheProAsnArg-----AsnSerAsnAspSerArgTyrProTyrThrAsnAsnProAsn 38
DB 43 ACAACCTTAACAGTGAATTCGAACCTAACCTAATCATATCTTTAGCTGACAACTCAAT 102
QY 39 GlnProLeuGlnAsnThrAsnTyrIlyGluTTrPLeuAsnMetCysGln---GlyAsnThr 57
DB 103 TCAACACCTAGGAAGAAATTAAATTAAAGAAATTTTAAAGATGACTGAAGACAGTTCTAG 162
QY 58 GlnTyrGlyAspAsnAspPheGluThrPheAlaSerAlaAspThrIleAla---AlaValSer 76
DB 163 GAAGTGTACACAAAC-----TCTACAGTAAAGATGACAGTTGGG 201
QY 77 AlaGlyThrIleValSerGlyThrLeuLeuAlaGlyIleGlyGlyLeuThrSerIleSer 96
DB 202 ACAGGAATTTCTGTTAGGGGAGATTTAGGTTGAGGA----- 243
QY 97 GlyProIleGlyIleIleGlyAlaIleIleIleSerPheGlyThrLeuIleThrValPhe 116
DB 244 -----GTTCCATTGGCTGGGCGACCTCACTTATTCATCATCTTCTTAACACTATA 297
QY 117 TrpProAlaGlyGluGlnAspLysThrValTTrpGlnPheIleLysMetGlyGluIle 136
DB 298 TGGCCACAGT-----GATGCTGACCCATGGAAGGCTTTATGTCACAACTGTAAGTA 348
QY 137 PheValAspThrProLeuThrGlnSerIleLysGlnLeuLysLeuGlnThrLeuGlnGly 156
DB 349 CTGATGTGATTAAGAAATTAAGAGAGATGCTAAAGAAATTAAGAGAGATGCTTCAAGTTAC 408
QY 157 PheArgGlnIleLeuGlnInsertYrAsnThrAlaLeuAspAspTrpArgLysLeuLysArg 176

```

```

DB 409 CTTCAAAATTAATTGGAAGATTAATGTTAATGCGTTAAATTCCTCGAAGAAA----- 459
QY 177 LeuGlnAlaProGlyLeuProProSerSerAlaLeuGlnGlnAlaAlaLeuThrLeuLys 196
DB 460 -----ACACCTTTAAGTTTGCGAAGTAAAGAAAGGCGAAGATCGAATTAAG 504
QY 197 IleArgPheGluAsnValHisAsnAspPheIleArgGluIleProGlyPheGlnLeuGln 216
DB 505 GAACCTTTTCTCAAGCAGAAAGATCTTTTGGTAATTCATTCATCCGCTCATTTGCACTTCC 564
QY 217 ThrTyrLysThrLeuLeuLeuProIleTyrAlaGlnAlaAlaAsnPheHisLeuAsnLeu 236
DB 555 AAATTCGAAGTCTGTTCTTACCAACATATGCAAGCTGCAAAATACATATTATTTGCTA 624
QY 237 LeuGlnGlnGlyAlaGluLeuAlaAspGluTTrpAsnAlaAspIleHisProSerGlnIle 256
DB 625 TTTAAAGATGCTCAAGTTTGGAGAAAGATGGGATATTC----- 666
QY 257 GluProAsnAlaGlyThrSerAspAspTyrTyrLysLeuLeuLysGluAsnIleProLys 276
DB 667 -----TCAGAAAGATGTTGCTGAATTTATTCATGACATTAACCTTACA---CAACA 717
QY 277 TyrSerAsnTyrCysAlaAsnThrTyrArgThrGlyLeuLysAsnLeuArgAspGluPro 296
DB 718 TACACTGACACATGTTGTTAATGATATATGTTAATGTTAATGTTAAGAGTTCAACT 777
QY 297 AsnMetLysTrpSerIlePheAsnAspTyrArgArgTyrMetThrIleThrValLeuAsp 316
DB 778 TATGATGATGAGGTGCMAATTTCACCGTTTTCGCAAGAAATGACTTAACTGATTAAGAT 837
QY 317 ThrIleSerGlnPheSerLeuLysAspIleLysArgTyrArgAspSerIleGlyIle 336
DB 838 CTAATGTGACTTTCCATTTATATATATGCGTTATACTCA----- 879
QY 337 GluValLysGlyIleLysAsnGluLeuThrArgGluIleTyrThrGluIleAsnPhe 356
DB 880 -----AAAGGGCTTAAACAGACTAACAGACATTTTACGATCCGATCTTTTCT 933
QY 357 AspArgLeuProGlnLeuArgValGlnProAsnLeuAlaThrMetGluTyrAsnLeuThr 376
DB 934 CTTAGACACCA---CTTGCGTACGAGCAACCTTTTTCAGATGATGAAGAAACTATTCGA 990
QY 377 ArgAlaSerPheLysLeuPheSerPheLeuGlnGlnPheIlePheTyrThr 393
DB 991 AAACCT-----CATTTATTTGATTAATTTACAGGGGATGGAATTCATACGGCTTCA 1044
QY 394 -----GluAsnThrAsnPhe-----GlyAsnArgLeuValGlyIle 405
DB 1045 CTTGCTTACTTTGGGAAAGATTTCTTCATTTATTTGCTGCTGTAATTA-----GTA 1095
QY 406 SerAsnArgAspAlaProThrTyrSerAsnThrIleThrGlnThrLeuTyrGlyGluArg 425
DB 1096 GAACCTAGACCTAGATAGATCTAGTAAGACAAATTAATCTCCCATTTTATGAGATTA 1155
QY 426 ThrGlySerProThrThrLysThrIleArgProPheGlnSer-----TyrLysVal 442
DB 1156 TCTACTGAACTGTGCAAAAGCTA-----ACCTTGATGAGCAAAAAGTTTATCAACT 1209
QY 443 SerIleValThrAspArgGlnSerPro-----ValSerPro 455
DB 1210 ATAGCTAATACAGAGTACCGGCTTGCCGAATGTAAGTATATTTAGGTGTACGAA 1269
QY 456 IleGlnProHisPheIleIleAsnGlnIleGluLeuTyrLeuAsnGlySerSerAsnAsn 475
DB 1270 GTTGAT-----TTTAGTCATATGATGATCAAAAAGAACTAGTACACAA 1317
QY 476 ThrLeuLysTyrSerAlaGlyLysSerLeuSerAsnTyrGlnAsnThrThrPhePheGln 495
DB 1318 ACATATGATTCAAAAGAAACAATGCGCATGTAAATGTCAGAGATTTCTATTAACCAATTA 1377
QY 496 PheProArgLysLysAspCysAsnLeuValIleAspProGlyCysSerProAsnPheAsn 515

```

1378 CCGCCAGAAACACAGAT-----GAACTGTGAAAA 1410
QY AsnTyrSerHisIleLeuSerHisPheSerLeuPheThrTyrSerTyrValIleGlyLeu 535
DB 1411 GCATATAGTCACACTTATTAACGGGAAATTTCTTA-----ATG 1452
QY 536 GlnLeuGlnIleLeuAspThrGlyValLeuGlyTyrThrHisSerSerValAspArgTyr 555
DB 1453 CAGGACCGTCGTGAACAAATTCATTTTCTTACCTTGACACATAGAGGTGACACTTTT 1512
QY AsnAlaIleSerAspLysIleIleThrMetIleProAlaIleLeuGlyAsnAsnLeuAsp 575
DB 1513 AATACAAATGATGCTGAAAAGATTAATCTCAAGTAGTAAAGCAATAGCCCTTGCT 1572
QY 576 ThrAsnSerLysValIleGlyGlyProGlyHisThrGlyGlyAsnLeuValTyrLeuGln 595
DB 1573 TCAGGTGCTTCATTAATTTGAAGTCCAGAGTTCACAGAGAAATTTACTATCTTAA 1632
QY 596 -----SerGlnGlyArgLeuGlnIleThrCysGluThrProAsnSerThrGln 611
DB 1633 GAATCTAGTAATTCATTCATTCATTAATTAAGTTACATTAATTCAGACGCTTGTACAA 1692
QY 612 SerTyrPheIleArgLeuAspGlyTyrAlaThrAsnGlyAlaGlyAsnThrLeuProAsnIle 631
DB 1693 CGATATCGGTAGAGATAGCTATGCTTCTACC-----ACTACCTTA 1734
QY 632 SerLeuThrIleProGlyValIleGlyIleProProGlnArgLeuAsnAsnThrPheSer 651
DB 1735 CGACTTTTGTG-----CAAAATTCAAACATGATTTTCTT 1770
QY 652 GlyThrAsnTyrAsn-----AsnLeuGlnTyrGlyAspPheGly 664
DB 1771 GTCATCTACATTAATAAATACTATGATTAAGATGATGATTAATTAACATCAACATTTGAT 1830
QY 665 TyrPheGlnPheProSerThrValThrIleuProLeuAsnArgAsnIleProPheIlePhe 684
DB 1831 CTCGCACTACTAATTTCTAATATGAGGCTTCTCGCGGTAGAAAGAACTTAATTAATGAGA 1890
QY 685 AsnArgAlaAspValSerAsnSerIleLeuIleIleAspLysIleGlyPheIleProIle 704
DB 1891 GCAAGATCTTTCGTTCTTAATGAAAAAATCTATATAGATAGATAGATTTATCCACAGTA 1950
RESULT 15
US-07-973-320-3
; Sequence 3, Application US/07973320
; Patent No. 5286486
; GENERAL INFORMATION:
; APPLICANT: Payne, Jewel M.
; APPLICANT: Fu, Jenny M.
; TITLE OF INVENTION: No. 5286486el Bacillus thuringiensis Gene
; TITLE OF INVENTION: Encoding a Coleopteran-Active Toxin
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: David R. Saliwanchik
; STREET: 2421 N.W. 41st Street, Suite A-1
; CITY: Gainesville
; STATE: FL
; COUNTRY: USA
; ZIP: 32606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/973,320
; FILING DATE: 19921106
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/788,638
; FILING DATE: 6-NOV-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Saliwanchik, David R.

REGISTRATION NUMBER: 31,794
REFERENCE/DOCKET NUMBER: MA68.C1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 904-375-8100
TELEFAX: 904-372-5800
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 3414 base pairs
TYPE: NUCLEIC ACID
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Bacillus thuringiensis
STRAIN: kumamotoensis
INDIVIDUAL ISOLATE: HD867
IMMEDIATE SOURCE:
LIBRARY: Lamdegem (TW)-11 library of J.M. Fu
CLONE: 867
US-07-973-320-3
Alignment Scores:
Pred. No.: 1.18e-59 Length: 3414
Score: 660.50 Matches: 223
Percent Similarity: 43.1% Conservative: 112
Best Local Similarity: 28.7% Mismatches: 286
Query Match: 17.1% Indels: 157
DB: 2 Gaps: 37
US-10-783-417-2 (1-735) x US-07-973-320-3 (1-3414)
QY 1 MetAsnGlnAsnAspAsnAsnGlnIleIleAspSerHisThrSerProTyr 20
DB 1 ATGAATTTAAATTAATTAGTGA--TATGAA-----GATAGT----- 36
QY 21 PheProAsnArgAsnSerAsnAspSer---ArgTyrProTyrThrAsnAsnProAsnGln 39
DB 37 -----AATAGAACCTTAATTAATTTCTCATATTAATCTTACGCAAAAAGCATATACACA 90
QY 40 ProLeuGlnAsnThrAsnTyrIleGlyIleuThrLeuAsnMetCysGlnGlyAsnThrGlnTyr 59
DB 91 TCATTAAGAATATGAACTACACAGAGATTTTATCTATATACAGAG----- 135
QY 60 GlyAspAsnPheGlnThrPheAlaSerAlaAspThrIleAlaValaSerAlaGlyThr 79
DB 136 AGGGAACAGCTGGAAGCACTGCTAGTGTATATACA-----GCTATTAAATGCTGACTT 189
QY 80 IleValSerGlyThrLeuLeuAlaGlyIleGlyIleuThrSerIleSerGlyProIle 99
DB 190 AGGTATACAGGGCTACACTAGTCCGTTAGT-----GTCCAGGT----- 231
QY 100 GlyIleIleGlyAlaIleIleIleSerPheGlyThrLeuIleThr---ValPheTyrPro 118
DB 232 -----GCAAGTTTATCATCACTTAATCTTACCTGAAAATTAACAGGCTTTATAGGCA 282
QY 119 AlaGlyGlnAspLysThrValTyrThrGlnPheIleLysMetGlyGlnIlePheVal 138
DB 283 -----CACGATAAATAATATTGGATGAATTTATAGACAGAAAGTAAACACTTATT 333
QY 139 AspThrProLeuThrGluSerIleLysGlnLeuLysLeuGlnThrLeuGlyGlyPheArg 158
DB 334 GAACAAAAAATAGAACATATATGCAAGAAATTAAGACACTTGCAGATTTAGAGGATTAAGA 393
QY 159 GlnIleLeuGlnSerTyrAsnThrAlaLeuAspAspTyrPargLysLeuLysArgLeuGln 178
DB 394 AATAACTTAAGATATATCAACAGCACTGAAGATTGG-----CTGAAC 438
QY 179 AlaProGlyLeuProProSerSerAlaLeuGlnGlnAlaIleuThrLeuLysIle--- 197
DB 439 AATCCTGATGATCCA-----GCAACTATTAACACAGATGATGAT 477

QY 198 ArgPheGluAsnValHisAsnAspPheIleArgGluIleProGlyPheGluLeuGluThr 217
 |||||
 |||||
 478 CGTTTCGTATATTAGATGCTTATTGATCATCATATGCGCATTTGAGTTGCTGG 537
 QY 218 TyrIleThrLeuLeuLeuProIleTyrAlaGlnAlaAlaAsnPheIleLeuLeuLeu 237
 |||||
 |||||
 538 TATGAATACCATTTACTTAACAGTTTACGACCAAGCGCAACCTTCATCTAGCTTATTA 597
 QY 238 GlnGlnGlyAlaGluLeuAlaAspGluTTPAsnAlaAspIleHisProSerGlnIleGln 257
 |||||
 |||||
 598 AGAGATTCTACTCTTTATGAGATTAATAGS-----GAATTCCTCAGAACCAATTTGAG 651
 QY 258 ProAsnAlaGlyThrSerAspAspTyrTyrIleLeuLeuLeuGluAsnIleProIleTyr 277
 |||||
 |||||
 652 GAAAT-----TATTAATCGTCAAAAGAAACATATTTCTGAATAT 690
 QY 278 SerAspTyrCysAlaAsnThrTyrArgThrGlyLeuLeuAsnLeuLeuArgAspGluProAsn 297
 |||||
 |||||
 691 TCTAACCATTCCTTAAAGTGTATTAATAGTGTCTTACGACATTTGAACGCTTCACTTAT 750
 QY 298 MetIleTyrSerIlePheAsnAspTyrArgArgTyrMetThrIleThrValIleAspThr 317
 |||||
 |||||
 751 GAACAAATGATTAATATATATCGTTTCGTAGAGAAATGATTAATATGATTAATAT 810
 QY 318 IleSerGlnPheSerLeuTyrAspIleLeuArgTyrArgAspSerIleGlyGlyIleGln 337
 |||||
 |||||
 811 GCGCTGATTTCTTATTTATGACCTCGAATGTAT-----TCAATGGAA 855
 QY 338 ValIleGlyIleIleAsnGluLeuThrArgGluIleTyrThrThrGluIleAsnPheAsp 357
 |||||
 |||||
 856 ACA-----AGTACGACGATTAAACGAGAGATGTATATACCATCAATTAGCTTGCTCA 906
 QY 358 ArgLeuProGlnLeuArgValGlnProAsnLeuAlaThrMetGluTyrAsnLeuThrArg 377
 |||||
 |||||
 907 ---ATTAGCAATCCAGGTATAGGTCCAGATTTCAGATGGAATAATCTGCGATTAA 963
 QY 378 AlaSerPheIleLeuPheSerPheLeuGluGlnPheIlePheTyrThrGluAsnThrAsn 397
 |||||
 |||||
 964 AACCCA---CAGCTTGTGATTTATTTAGATGAGCTTATATATATACA----- 1008
 QY 398 PheGlyAsnArgLeuValGlyIleSerAsnArgAspAlaPro----- 411
 |||||
 |||||
 1009 -----TCAAAATATTAAGCATTTTCAACATGAGATTCACACGACCTATTTATGAGAT 1062
 QY 412 -----ThrTyrSerAsnThrIleThrGluThrLeuTyr 422
 |||||
 |||||
 1063 GCACATTAAGTTAGCTTAAACAATCGAGCAATGCAATTATATATACACAGGCATATAT 1122
 QY 423 GlyGluArgThrGly-----SerProThrThrIleArgProPheGluSerTyr 440
 |||||
 |||||
 1123 GGTAAACAAGTGTATATTTATTCATCGAGGCAATATTCATTTAGAGGTAAATGATATCAT 1182
 QY 441 LysVal-----SerIleValThrAspArgGlnSerProValSerProIle 456
 |||||
 |||||
 1183 AGAACATTAGCAGCTCCATCAGTTGA-----GTTTATCCGTAT 1221
 QY 457 GlnProHisPheIleIleAsnGlnIleGluLeuTyr-----Leu 469
 |||||
 |||||
 1222 ACTCAAGATTAATGTGTGCGAGAGTTGAGTTTACGCTGTAAAGGCAAGCATATAT 1281
 QY 470 AsnGlySerSerAsnAsnThrIleLeuTyrTyrSerAlaGlyGlySerLeuSerAsnTyrGln 489
 |||||
 |||||
 1282 AGAGAGATTAACAATATGATCTGACGTATGAT----- 1314
 QY 490 AsnThrThrPhePheGlnPheProArgLysLysAspCysAsnLeuValIleAspProGly 509
 |||||
 |||||
 1315 -----TCTATGATCAATTACCC-----CCAGAC 1338
 QY 510 CysSerProAsnAsnAsnThrTyrSerHisIleLeuSerHis-----PheSerLeu 526
 |||||
 |||||
 1339 GGAGAACCAATACACAAAAATACACTCATGATTAATGTCATGCTACAGCTATATCTAAA 1398
 QY 527 PheThrTyrSerTyrValIleGlyLeuGlnLeuGlnIleLeuAspThrGly----- 543

Db 1399 TCAACTCCGGAATAT-----GATTAATGCTACTATCCCG 1431
 QY 544 ValIleGlyTyrThrHisSerSerValAspArgTyrAsnAlaIleSerAspLysIleIle 563
 |||||
 |||||
 1432 ATCTTTTCTTGACGATTAAGATGCGAGTATTAACAATGAATCTATCCAAACAAATTC 1491
 QY 564 ThrMetIleProAlaIleLeuGlyAsnAsnLeuAspThrAsnSerLysValIleGlyGly 583
 |||||
 |||||
 1492 ACBAATTTCCAGCTGTAAATATGTAATACATGATGATCAATCTACATTTGCAACAGG 1551
 QY 584 ProGlyHisThrGlyValAsnLeuValTyrLeuGlnSerGlnGlyValArgLeuGluIleThr 603
 |||||
 |||||
 1552 CCTGATTTACAGGTGAGATTTAGTTAAGACAGGAGATTAATGTTATATAGAGATATA 1611
 QY 604 CysGluThrProAsnSer-----ThrGlnSerTyrPheIleArgLeuArgTyrAlaThr 621
 |||||
 |||||
 1612 AAGGCTACCGTAACTCACACCTTCTCAAAATTTATCGTTAGAGCTTGATACGCCACT 1671
 QY 622 AsnGlyAlaGly-----AsnThrLeuProAsnIleSerLeuThrIleProGlyValIle 639
 |||||
 |||||
 1672 AATGTTTCTGACATTCACGCTATATTAATGATTAATAATACGCTT----- 1719
 QY 640 GlyIleProProGlnArgLeuAsnAsnThrPheSerGlyThrAsnTyr-----AsnAsnLeu 658
 |||||
 |||||
 1720 -----CAAAAGAAAGTTTCAAAATACTGTAGAAACAATAGGTGAAAGAAAGATTAA 1770
 QY 659 GlnTyrGlyAspPheGlyTyrPheGlnPheProSerThrValThrLeuProLeuAsnArg 678
 |||||
 |||||
 1771 ACCATGTTCTCATTTGAGATTAATGAAATTTCTACACACCATTCATTTCCG--GATTAAG 1827
 QY 679 AsnIleProPheIlePheAsnArgAlaAspValSer---AsnSerIleLeuIleIleAsp 697
 |||||
 |||||
 1828 CATCCAAAATACACTCTTCATTTAAGATTAATGATTAAGATTAATCAATTTATATGAT 1887
 QY 698 LysIleGluPheIleProIleThrSerSerMetHisGlnAsnAspGlyLysGlnLysLeu 717
 |||||
 |||||
 1888 TCAATGGAATTTATCCCTGTAGATTAATATGAT-----GAAAAAGAAAAACATA 1938
 QY 718 GluThrIleGlnThrLysIleAsnThrPhePheThrAsnHisThrLysThrLeu 735
 |||||
 |||||
 1939 GAAAAAGCACAGAAAGCCGTGAATACCTTGTTTACAGAGGAAAGAAATGCACCTC 1992
 Db

Search completed: January 20, 2006, 22:34:14
 Job time : 411 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: January 20, 2006, 16:46:39 ; Search time 7267 Seconds
(without alignments)
4732.146 Million cell updates/sec

Title: US-10-783-417-2

Perfect score: 3869
Sequence: 1 MNQNNNDNNEYELIDSHTSPLY.....KLETIQKINTEFFTHNTKTL 735

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 41078325 seqs, 23393541228 residues
Total number of hits satisfying chosen parameters: 82156650

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODE=frame_p2n.model -DEV=xlh
-O=/cgn2.1/USPTO.spool/US10783417/runat.20012006.095010.23879/app.query.fasta_1.903
-DB=EST -QFMT=fastlap -SUFFIX=p2n.rst -MINMATCH=0.1 -LOOPEXT=0 -LIST=45
-UNITS=bits -START=1 -END=-1 -MATRIX=blonum62 -TRANS=human40.cdi -LISF=45
-DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pro -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US10783417@cgn2.1.5315@runat.20012006.095010.23879 -NCPU=6 -ICPU=3
-NO MMAP -NEG SCORES=0 -WAIT -DSPBLOCK=100 -DEV TIMEOUT=120 -WARN TIMEOUT=30
-THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7 -YGAPOP=10
-YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

EST:*
1: gb_esc1:*
2: gb_esc2:*
3: gb_esc3:*
4: gb_hic:*
5: gb_esc4:*
6: gb_esc5:*
7: gb_esc6:*
8: gb_esc7:*
9: gb_gsa1:*
10: gb_gsa2:*
11: gb_gsa3:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	147.5	3.8	6567	11	DD031803	DD031803 Homo sapi
2	124	3.2	13107	11	DD035722	DD035722 Homo sapi
3	121.5	3.1	3477	10	CL977084	CL977084 OlfCC029
4	121	3.1	1754	10	AG392933	AG392933 Mus muscu
5	119.5	3.1	2032	4	CNS0A8VA	BX819057 Arabidops
6	119.5	3.1	6777	10	CL967047	CL967047 OlfCC014
7	117.5	3.0	1986	10	AY416726	AY416726 Mus muscu

8	117	3.0	2157	4	AK089255	AK089255 Mus muscu
9	115	3.0	1178	1	AJ538005	AJ538005 OlfCC004
10	115	3.0	2322	1	CL960568	CL960568 OlfCC004
11	115	3.0	2460	4	CNS09YF2	BX822541 Arabidops
12	114.5	3.0	820	7	CF951415	CF951415 Ut-M-HL0-
13	112.5	2.9	1123	10	AG378595	AG378595 Mus muscu
14	112.5	2.9	2841	10	CL974457	CL974457 OlfCC025
15	111	2.9	1143	4	CNS0AJJE	BX829716 Arabidops
16	111	2.9	1616	10	AG396120	AG396120 Mus muscu
17	110.5	2.9	1105	3	BM463012	BM463012 ACENECOURT
18	109.5	2.8	752	7	CO096813	CO096813 GR_Ea20E
19	109.5	2.8	861	7	CO070923	CO070923 GR_Ba28D
20	109	2.8	1878	2	BF528790	BF528790 602041401
21	108.5	2.8	949	7	CV674212	CV674212 RET764_30
22	108.5	2.8	956	4	AY812233	AY812233 Schistoso
23	108.5	2.8	2282	4	AK044734	AK044734 Mus muscu
24	108.5	2.8	5001	4	CR860285	CR860285 Pongo pyg
25	107.5	2.8	2223	10	CL981417	CL981417 OlfCC045
26	107.5	2.8	2439	10	CL982610	CL982610 OlfCC048
27	107.5	2.8	2955	10	CL947058	CL947058 OlfCC002
28	107	2.8	859	9	AZ548115	AZ548115 EMTDL91TR
29	106.5	2.8	1541	4	BC015498	BC015498 Homo sapi
30	106.5	2.8	1598	4	BC030976	BC030976 Homo sapi
31	106.5	2.8	2355	10	AY413036	AY413036 Mus muscu
32	106.5	2.8	2948	4	CNS0A4RI	BX822972 Arabidops
33	106.5	2.8	4751	4	CR859511	CR859511 Pongo pyg
34	106	2.7	881	9	BH162188	BH162188 EMTYF54TR
35	106	2.7	910	9	AZ546690	AZ546690 ENTEN48TP
36	106	2.7	2520	10	AG280133	AG280133 Mus muscu
37	106	2.7	3054	10	CL974410	CL974410 OlfCC025
38	105.5	2.7	1092	10	CL065922	CL065922 CH216-105
39	105	2.7	622	7	CV176047	CV176047 SPBOROMA
40	105	2.7	697	7	CV533100	CV533100 LVS_035_B
41	105	2.7	747	3	BJ311235	BJ311235 Buj11235
42	105	2.7	16387	11	DQ038517	DQ038517 Pan trogl
43	105	2.7	16425	11	DQ038516	DQ038516 Homo sapi
44	104.5	2.7	458	1	AU006750	AU006750 AU006750
45	104.5	2.7	872	6	CD792427	CD792427 EST663788

ALIGNMENTS

RESULT 1	DD031803	6567 bp	DNA	linear	GSS 02-JUN-2005
LOCUS	DD031803				
DEFINITION	Homo sapiens HC2093 gene, VIRTUAL TRANSCRIPT, partial sequence,				
ACCESSION	DD031803				
VERSION	DD031803.1	GI:66683012			
KEYWORDS	GSS.				
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.				
AUTHORS	1 (baaes 1 to 6567)				
TITLE	Nielsen,R., Bustamante,C., Clark,A.G., Gianowski,S., Sackton,T.B., Hubisz,M.J., Fledel-Alon,A., Tanenbaum,D.M., Civeleo,D., White,T.J., Shinsky,J.J., Adams,M.D. and Cargill,M.				
JOURNAL	A Scan for Positively Selected Genes in the Genomes of Humans and Chimpanzees				
PUBMED	(et) PLOS Biol. 3 (6), E170 (2005)				
REFERENCES	2 (baaes 1 to 6567)				
AUTHORS	Nielsen,R., Bustamante,C., Clark,A.G., Gianowski,S., Sackton,T.B., Hubisz,M.J., Fledel-Alon,A., Tanenbaum,D.M., Civeleo,D., White,T.J., Shinsky,J.J., Adams,M.D. and Cargill,M.				
TITLE	Direct Substitution				
COMMENT	Submitted (05-MAY-2005) Celera Genomics, 45 West Gude Drive, Rockville, MD 20850, USA This sequence was made by sequencing genomic exons and ordering them based on alignment. Translation starts at the beginning of alignment.				

FEATURES
source

Location/Qualifiers
1. 6567
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="12"
<1..>6567
/locus_tag="HC2093"

ORIGIN

Alignment Scores:

Pred. No.: 0.000318 Length: 6567
Score: 247.50 Matches: 160
Percent Similarity: 33.6% Conservative: 101
Best Local Similarity: 20.6% Mismatches: 238
Query Match: 3.8% Indels: 277
DB: 11 Gaps: 40

US-10-783-417-2 (1-735) x DQ031803 (1-6567)

QY 10 TyrGluIleIleAspSerHisThrSerPro---TyrPheProAsnArgAsnSerAsnAsp 28
DB 1951 TATGAGAGCTATATAAAATATAGTACTTATATATAGAACACATCAACACAGAC 2010
QY 29 -----SerArgTyrProTyrThrAsnAsnProAsnGlnProLeuGlnAsnThr 44
DB 2011 ATATATATTAGAGACTTATAGACCTCACACC-----CTCTATTAACATT 2052
QY 45 AsnTyrLysGluThrLeuAsnMetCysGlnGlyAsnThrGlnTyrGlyAspAsnHeglu 64
DB 2053 TCTGTAGGCTTACACCCAGATTGGTCAATGCGCAATCAAGTA-----TCTTCTTACTC 2106
QY 65 ThrPheAlaSerAlaAspThrIleAlaAlaValaSerAlaGlyThrIleValSerGlyThr 84
DB 2107 TCTGTAGAGACTTCGAGACTGTGCGCTGATAGTGCACGAAATATACCTTACAAAAT 2166
QY 85 LeuLeuAlaGlyIleGlyGlyLeuThrSerIle-----SerGlyProIleGlyIleIle 102
DB 2167 ATTCTTCTCGAGAGATTGAGCTATCATCTCTCCCAAGTAAGTCCCAATGAAATCATTA 2226
QY 103 GlyAlaIleIleIleSerPheGlyThrLeuIleThrValPheThrProAlaGlyGluGln 122
DB 2227 AAAAAATATACAAATTATCTCAAGAGA-----AGTAATGAAATATAG 2268
QY 123 AspLysThrValTyrThrGlnPheIleLeuMetGlyGluIlePheValAspThrProLeu 142
DB 2269 GAAAGACCTATAAATACA-----ACCTCTTAA 2295
QY 143 ThrGluSerIleLeuGlnLeuLysLeuGlnThrLeu----- 154
DB 2296 ACCCAAAACATTAAAGTACTGAAAGAAATATACCAATATATATGAGGTCTGCTAGT 2355
QY 155 -----GluGlyPheArgGlnIleLeuGlnSerTyrAsnThrAlaLeuAspAsp 170
DB 2356 ACACTGAAAGGTGAAAGAGTTCGAGAGTCTCCCATTAAGTATATCTGACGAGAAAT--- 2412
QY 171 TrpArgLysLeuLysArgLeuGlnAlaProGlyLeuProPheSer-----SerAlaLeu 188
DB 2413 -----GCTCTGATTTCTCCCTCCAGACTTCTCTGTAAAA 2448
QY 189 GlnGlnAlaAlaLeuThrLeuLysIleArgPheGluAsnValHisAsnAspPheIleArg 208
DB 2449 CAGTGTCTGTGTCTACCGGTGAAAGTTGTCATGGCAA----- 2484
QY 209 GluIleProGlyPheGln-----LeuGluThrTyrLysThrLeuLeuProIle 225
DB 2485 -----CCACCCCTGAGGCAAAATGAAATATCTTTATTAACAGTTTAT-----GTC 2532
QY 226 TyrAlaGlnAlaAlaAsnPheHisLeuAsnLeuLeuGlnGlyAlaGlyLeuAlaAsp 245
DB 2533 TGGATTAATCATCTATAAACTATTAATGTCACTGAAACATCATTTGAGTTATCAAGT 2592
QY 246 -----GluThrAsnAlaAspIleHisProSer----- 254

DB 2593 TTGATTATTAATGTTGAATACAGTCTTATATGAAACAGCTAGACCAAGATTGGTATGG 2652
QY 255 -----GlnIleGluProAsnAlaGlyThr 262
DB 2653 AAAACAAAGACAAATATCATTTAGCTTTCAACACCAAGGAGGACCAAGCATCTCTCC 2712
QY 263 SerAspAspTyrTyrLysLeuLeuLysGluAsn----- 273
DB 2713 AAAGATGTTTATATGAAACCTCAGTTCTTCATCAATTAATCTTTTTCGACACCTCCT 2772
QY 274 -----IleProLysTyrSerAsnTyrCysAlaAsnThrTyrArgThr 287
DB 2773 TCAAAACCTTAATGGGATTATACAAATATTAATCTGTATTATTAAGAAATATCTCAGTACT 2832
QY 288 GlyLeuLysAsnLeuArgAspGluProAsnMetLysTrpSerIlePheAsnAspTyrArg 307
DB 2833 TTTATGCAAGNNNNNNNNNTCCAT-----GAAGTAACCAATGACTTTGAC 2877
QY 308 ArgTyrMetThrIleThrValLeuAspThrIleSerGlnPheSerLeuTyrAspIle--- 326
DB 2878 AATATGACTGATATCCACAATTATATAGTAACTGACAAATATTCAGCTACTATCATTTGG 2937
QY 327 LysArgTyrArgAspSerIleGlyGly-----IleGluVal--- 338
DB 2938 TTAAACAGCAAGTACTTCAAGTGGAAATGGGAATAAACAGTGCATCATTAAGATATAC 2997
QY 339 -----LysGlyIleLysAsnGluLeuThrArgGlu---IleTyrThr 351
DB 2998 ACAAGTCAAGACATACCTGAAAGGTTTGTGGAAACCTGACTTACGAATCCATTTCGTCA 3057
QY 352 ThrGluIleAsnPheAspArgLeuProGlnLeuArg----- 363
DB 3058 ACTGCATTAATATGTAAGCTGGGTCCACCGGCTCAACCAACGGTCTAGTTCTTACTAT 3117
QY 364 -----ValGlnProAsnLeuAlaThrMetGlu 372
DB 3118 GTTTCACGTATCTTACACAGACTCTCGCCATGGAAGCACCTTGTTCATATATGAG 3177
QY 373 TyrAsnLeuThrArgAlaSerPheLysLeuPheSerPheLeuGluGlnPheIlePheTyr 392
DB 3178 AGAGCATTA-----TATTTGATTAATCTGGAAATATACCTGATTAAT 3219
QY 393 -----ThrGluAsnThrAsnPheGlyAsnArgLeuValGlyIleSerAsnArg 408
DB 3220 ATATTAATAAATTACTCCATCAACAGAAAGGA----- 3252
QY 409 AspAlaProThrTyrSerAsnThrIleThrGluThrLeuTyrGlyGluArgThrGlySer 428
DB 3253 -----TTCCTGATATCTTATCTGCCCAAGCTATAC----- 3282
QY 429 ProThrThrLysThrIleArgProPheGluSerTyrLysValSerIleValThrAspArg 448
DB 3283 -----ATCAAGACTGAAGAA 3297
QY 449 GlnSerProProValSerProIleGlnProHisPheIleIleAsnGlnIleGluLeuTyr 468
DB 3298 GATGCCCAAGAACTTACCAATA-----ATCAACACTTT 3333
QY 469 LeuAsnGlySerSerAsnAsnThrLeu-----LysTyrSerAlaGlyGly 483
DB 3334 AAAAACTTTCTCTTACCTCAGTTCTCTTATCATGAGATCCCAAGTAAGCAATGGT 3393
QY 484 SerLeuSerAsnTyrGluAsnThrThrPhePheGlnIleProArgLysLeuAspCysAsn 503
DB 3394 GCATTAATTAAGTTATGATTTTAACT-----TTACACGACCAAAATGAA----- 3435
QY 504 LeuValIleAspProGlyCysSerProAsnPheAsnAsnTyrSerHisIle----- 520
DB 3436 -----AATTAATCTTCACTTACTTCTGAT 3459
QY 521 -----LeuSerHisPheSerLeuPheThrTyrSerTyrValIle 533

Db 3460 AATTACATATATTTGGAGAGCTTCCACATTATCATTTATAGCTTTTTCCTCCGCA 3519
 QY 534 GYLeuGlnLeuGlnIleLeuAspThrGlyValIeu---GlyThrPhthSerSerVal 552
 Db 3520 AGAAGTGAAGAAAGACTGCTGCTTCAGTATCTTTTTCATCAGATGAGTCAGTG 3579
 QY 553 -----AaPArgTyAsnAlaIleSerAspLysIle----- 562
 Db 3580 CCGTATAGCAGCTCCACAAAATTTGACTTATATCACTGACTTACAGCTTTGATGCGTG 3639
 QY 563 -----IleThMetIleProAlaIleLysGlyAsn 572
 Db 3640 AATAGAGCCCAAGTCTCTTCAGAGTGGTATGTAAAGTATATAGTTTAAATTCAT 3699
 QY 573 AsnLeuAspThrAsnSer-----LysValIleGluGlyProGlyHisThrGlyGly 589
 Db 3700 GAACATGAACATGACCTATATATATTAAGAAATATATACAGA---TTTAAACTGAAGCC 3756
 QY 590 AsnLeuValTyLeuGlnSerGlnGlyArgLeuGluIleThrCysGluThrProAsnSer 609
 Db 3757 AACTGTGTGACTGAA-----CCAGTCAGC 3783
 QY 610 ThrGlnSerTyPheIleArgLeuArgTyAlaThrAsn---GlyAlaGlyAsnThrLeu 628
 Db 3784 ACC-----TACTCTATCCGTATCTGCGTTCACCAAGTTGGAATGCAATCAATT 3837
 QY 629 ProAsnIle-----SerLeuThrIleProGlyValIle 639
 Db 3838 ACTAATGTATGTAATAATTCAACCCCAAGATCAGTTCAGATGTCGTG 3885
 RESULT 2
 LOCUS DQ035722 13107 bp DNA linear GSS 02-JUN-2005
 DEFINITION Homo sapiens DNAs gene, VIRUTAL TRANSCRIPT, partial sequence.
 ACCESSION DQ035722
 VERSION DQ035722.1 GI:66886931
 KEYWORDS GSS.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
 Homiidae; Homo.
 REFERENCE 1 (bases 1 to 13107)
 AUTHORS Nielsen,R., Bustamante,C., Clark,A.G., Glanowski,S., Sackton,T.B.,
 Hubisz,M.J., Fledel-Alon,A., Tanenbaum,D.M., Civeille,D.,
 White,T.J., Snieky,J.J., Adams,M.D. and Cargill,M.
 TITLE A Scan for Positively Selected Genes in the Genomes of Humans and
 Chimpanzees
 JOURNAL (et) PLOS Biol. 3 (6), E170 (2005)
 PUBMED 15869325
 REFERENCE 2 (bases 1 to 13107)
 AUTHORS Nielsen,R., Bustamante,C., Clark,A.G., Glanowski,S., Sackton,T.B.,
 Hubisz,M.J., Fledel-Alon,A., Tanenbaum,D.M., Civeille,D.,
 White,T.J., Snieky,J.J., Adams,M.D. and Cargill,M.
 TITLE Direct Submission
 JOURNAL Submitted (05-MAY-2005) Celera Genomics, 45 West Gude Drive,
 Rockville, MD 20850, USA
 COMMENT This sequence was made by sequencing genomic exons and ordering
 them based on alignment. Translation starts at the beginning of
 alignment.
 FEATURES
 source
 1..13107
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 /chromosome="5"
 <1..>13107
 /gene="DNAs"
 /locus_tag="HC4910"
 ORIGIN
 Alignment Scores:

Pred. No.: 0.536 Length: 13107
 Score: 124.00 Matches: 140
 Percent Similarity: 35.2% Conservative: 86
 Best Local Similarity: 21.8% Mismatches: 218
 Query Match: 3.2% Indels: 198
 Gaps: 35
 US-10-783-417-2 (1-735) x DQ035722 (1-13107)
 QY 97 GlyProIleGlyIleIleGlyAlaIleIleSerPheGlyThrLeuIleThrValPhe 116
 Db 3676 GCACCTATTAGAGAAATCTTATGCCCCGCTTAAACAGATATGAGACTTCGTAGACA----- 3729
 QY 117 TrpProIaGlyGluAsnAspLysThr-----ValThrPheGlnPheIle 131
 Db 3730 -----AGGAAAGAGATGACAAAGTTGATACAGTCACTGACTGAGCCGTTGAGAAAGCG 3783
 QY 132 ---LysMetGlyGluIlePheValAsp---ThrProLeuThrGluSerIleLysGlnLeu 149
 Db 3784 GCACGTGCTGGGAGAGCTCCAGAAATTAATGATCTACCTGACGCCCTTCAAGAAAG 3843
 QY 150 LysLeuGlnThrLeuGlnGlyPheArgGlnIleLeuGlnSerTyAsnThrAlaLeuAsp 169
 Db 3844 CTTATATTAGTGTGAGAGGTATCTCCCAAGATTGTCCACAGATTATCTGACTATGAT 3903
 QY 170 AspTrpArgLysLeuLysArgLeuGlnAlaProGlyLeuProSerSerAlaLeuGln 189
 Db 3904 -----TTGAATGCTCCAAATGCTACCGCTTAAAGCCAGAAAGCC---AGT 3948
 QY 190 GlnAlaAlaLeuThrLeuLysIleArgPheGluAsnValHisAsnAspPheIleArgIle 209
 Db 3949 GACAGGCTTATCATGTTTTCAGAAATCATTTGATTAATATCTATCCGAAATACATCATAT 4008
 QY 210 IleProGlyPheGlnLeu-----GluThrTyLeuThrLeuLeuProIle 225
 Db 4009 ACTGAGAGAGAGAGAGCTTTTGGCTGCCAGCTACACAGTATCTCAGCTTCTTGAATA 4068
 QY 226 TyralaGlnAlaAlaAsnPheHisLeuAsnLeuGlnGlnGlyAlaGluLeuAlaAsp 245
 Db 4069 AAGAGCA-----CTAAATCTTACACAAAATATATATCTG----- 4107
 QY 246 GluTrpAsnAlaAspIleHisProSerGlnIleGluProAsnAlaGlyThrSerAspAsp 265
 Db 4108 -----TACACAGCTGTACAGAA-----ACTGTAAATACG 4137
 QY 266 TyrTyLeuLeuLeu-----LysGluAsnIleProLysTySerAsnTyCys----- 281
 Db 4138 TATTATGATATCTTTGTCAGAGGTGAATATTGAAAAAATTAACAATGAATCACTTAGAA 4197
 QY 282 ---AlaAsnThrTyArgThrGlyLeuLysAsnLeuAspGluProAsnMetLysTrp 300
 Db 4198 TTCAGAAACAGATGCGAAAGCTTCCCGGCTTGAAGAC-----TGG 4242
 QY 301 SerIlePheAsnAspTyArgArg-----308
 Db 4243 CAGGCTTTTGGACCTGAAGAAATCATGATGATTTAGCGAGCTTCCCGCTGCGTG 4302
 QY 309 ---TyMetThr-----IleThrValLeuAsp 316
 Db 4303 GAATCATGCGCAGTAAAGCCATGATGAGCGGCACTGGGAAAGATTAACACCTTCACC 4362
 QY 317 ThrIleSer-----GlnPheSerLeuTyAspIleLys----- 327
 Db 4363 GGGCAGACGTGATGATGGGGAATGAAGAACTTTAAGTTAAGAAATTCATGAGGACCT 4422
 QY 328 -----ArgTyArgAspSerIleGlyIleGluValIleGlyIleLys----- 342
 Db 4423 CTTCTGAATATTAAGAGAAATTAAGACACTCTGATACAGTCCGCTGAAGAGAGAGAC 4482
 QY 343 -----AenGlu----- 344
 Db 4483 ATTGAGCAAAAGCTGAAGCAAGTGAATTAATGATGAGCAATATTAACATTCACCTTCGCG 4542

```

Oy 345 -----LeuThrArgGluIleTyrThrThrGluIleAsnPro 356
Db 4543 AGCTTTAAACCCGTCGAGAGCTCTTGAAGAGAGACAGATCCGGAATCATCCGC 4602
Oy 357 AspArgLeuProGlnLeuArgValGlnProAsnLeuAlaThrMetGlnTyrAsnThr 376
Db 4603 AACATGAGAGACAGCTGATGTTGCTGGAGATCCCTACGACACAGGACAAATG--- 4659
Oy 377 ArgAlaSerPheLeuPheSerPheLeuGlnIlePheTyrThrGluAsn--- 395
Db 4660 -----CCATTCAA-----GCCAGATCAAAAATGGTGTCAGTACCTTCCAACTCA 4707
Oy 396 ThrAsnPheGlyAsnArgLeuValGlyIleSerAsn-----ArgAspAlaPro 411
Db 4708 ACAGACATCATCGAGAGCTGATGACGGTCAAAAACCTGTGATTTATTTAGAACCTGC 4767
Oy 412 ThrTyrSerAsnThrIleThrGluThrLeuTyrGlyGlnArgThrGlySerProThrThr 431
Db 4768 TTTGTGGAGAGAGACATTCGCCAAGCAGCTGCCCAAGAA-----GCC 4809
Oy 432 LysThrIleArgProPheGluSerTyrLysValSerIleValThrAspArgGlnSerPro 451
Db 4810 AAGCGCTTTCTTAACATAGATAAATCTTGGGTGAAGATCATCATCGGACATGAATG 4869
Oy 452 Pro-----ValSerProIleGlnProHisPhe 460
Db 4870 CCCAGTATGATCGACAGCTGCTGTGGAGATGAGACCTCGGCGGACAGCTGTACCAAC- 4926
Oy 461 IleIleAsnGlnIleGluLeuTyrIleuAsnGlySerSerAsnAsnThrLeuLysTyrSer 480
Db 4927 TTGCTGACACAGTTGGAATA----- 4947
Oy 481 AlaGlyLysSerLeuSerAsnTyrGlnAsnThrThrPhePheGlnPheProArgLysLys 500
Db 4948 TGCCGAAATATCCCTTACTCGGTACTTGAGAAAAAACACATGCTTCTCTCGG----- 5001
Oy 501 AspCysAsnLeuValIleAspProGly-----CysSerProAsn 513
Db 5002 ---TTTTCTTCTGCTCAGATCTCGCCCTTACAGATTTCTGGGACAGCGTCGACATCC 5058
Oy 514 PheAsnAsnTyrSerHisIleLeuSerHisPheSerLeuPheThrTyrSerTyrValIle 533
Db 5059 CACACTATACAGGCCCATTTGCTGAATGTGTGAC----- 5094
Oy 534 GlyLeuGlnLeuGlnIleLeuAspThrGlyValLeuGlyTyrThrHisSerSerValAsp 553
Db 5095 -----AACATTAATCTGTCAAG 5112
Oy 554 ArgTyrAsnAlaIleSerAspLysIleIleThrMetIleProAlaIleLysGlyAsnAsn 573
Db 5113 TTCCACGAAAGATCATGATCGAAATTTCTGCA---ATTCTCTCAAGAGGTGAGAG 5169
Oy 574 LeuAspThrAsnSerLysValIleGluGlyProGlyHisThrGlyIleAsnLeu---Val 592
Db 5170 ATTGAA-----TTGATTAACCTGTGCATCGCAGAGGCGCATGTGAAGT 5214
Oy 593 TyrLeuGlnSerGln--GlyArgLeuGluIle-----ThrCysGlnThrProAsn 609
Db 5215 TGCGTTAATCTCTTTGGAAGAAATCTCAGTCTCATTCATGATCTGTGTCGACGCA 5274
Oy 609 eThrGlnSerTyrPheIleArgLeuArgTyrAlaThrAsnGlyAlaGlyAsnThrLeuP 629
Db 5275 GCCCGAAA-----TATTCAGAAACAGGTTTCCAACTACGAAATTTCTTCACTTC 5328
Oy 629 ro 629
Db 5329 CC 5330

```

```

RESULT 3
LOCUS CL977084 3477 bp DNA linear GSS 21-SEP-2004
DEFINITION OslPCC029775 Oryza sativa Expressed Library Oryza sativa (indica
cultivar-group) genomic, genomic survey sequence.

```

```

ACCESSION CL977084
VERSION CL977084.1 GI:52408678
KEYWORDS GSS.
SOURCE Oryza sativa (indica cultivar-group)
ORGANISM Oryza sativa (indica cultivar-group)
Bukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzaceae; Oryza.
1 (bases 1 to 3477)
Me,L., Wang,J., Chen,C., Liu,X., Su,N., Li,L., Wang,X., Cao,M.,
Jiao,Y., Sun,N., Zhang,X., Bao,J., Sun,D., Zhao,H., Yuan,L.,
Wong,G.K.S., Deng,X.W. and Wang,J.
An analysis of transcriptional regulation of the rice genome and
its comparison to Arabidopsis
Unpublished (2004)
JOURNAL
COMMENT
Contact: Chen Chen
Department of Bioinformatic
Beijing Institute of Genomics
Chinese Academy of Sciences, Beijing 101300, China
Tel: 86-10-80481559
Fax: 86-10-80488676
Email: chenchen@genomics.org.cn
Rice genomic sequence.
Class: exon-trapped.
Location/Qualifiers
source 1..3477
/organism="Oryza sativa (indica cultivar-group)"
/mol_type="genomic DNA"
/db_xref="taxon:39946"
/clone_lib="Oryza sativa Expressed Library"
/note="Oryza sativa exon trapped genomic sequences "

```

ORIGIN

Alignment Scores:

Pred. No.:	0.122	Length:	3477
Score:	121.50	Matches:	144
Percent Similarity:	31.84	Conservative:	79
Best Local Similarity:	20.54	Mismatches:	243
Query Match:	3.14	Indels:	235
DB:	10	Gaps:	31

US-10-783-417-2 (1-735) x CL977084 (1-3477)

```

Oy 132 LysMetGlyIleIlePheValAsp---ThrProLeuThrGluSerIleLys----- 147
Db 1105 AAAATGACGAACCTTACTACAGCAAAACCAATTCAGCGTTCATCCCAAGAAATA 1164
Oy 148 -----GlnLeuLysLeuGlnThrLeuGlnGlyPheArg 158
Db 1165 GGAATGCTAGCCAACTGCAAGTTTGCAGCTAAGCAACAACACCTTGAGCGG----- 1218
Oy 159 GlnIleLeuGlnSerTyrAsnThrAlaLeuAspAspTPrAspGlyLeuLysArgLeuGln 178
Db 1219 -----GAGATACCCACAGCTCTAGCAAAATCTGACCAACCTAGCTACTTGA 1266
Oy 179 AlaProGlyLeuProProSerSerAlaLeuGlnGlnAlaAlaLeuThrLeu---LysIle 197
Db 1267 CTATATGGTATATGAATCTGCGGCGCTATATCCCAAAACCTGTGACAGCTACCAAGAG 1326
Oy 198 ArgPheGluAsnVal---HisAsnAspPheIleArgGluIleProGlyPheGlnLeuGln 216
Db 1327 CAATATCTTAGCTTAGCAAAAACAACTTACGCGCAAAATCCGCGGTATATCAAT 1386
Oy 217 ThrTyrThrThrLeuLeuProLeyTyrAlaGlnAlaAlaAsnPheHisLeuAsnLeu 236
Db 1387 CTAAACAAAGTGAGAAAACCTTACTATACAAAACAA----- 1425
Oy 237 LeuGlnGlnGlyAlaGluLeuAlaAspGluTrpAsnAlaAspIleHisProSerGlnIle 256
Db 1425 ----- 1425
Oy 257 GluProAsnAlaGlyThrSerAspAspTyrTyrLysLeuLeuLysGluAsnIleProLys 276

```

1426 -----GTCCAGGGTTCATCCCAAA 1446
 QY 277 TyrSerAsnTyrCysAlaAsnThrTyrArgThrGlyLeuIleuAsnLeuArgAspGluPro 296
 DB 1447 GAATAGGAATGCTTCCCACTGCAGCTTTAGGCTAGGCAAAATACCTTGAACGA 1506
 QY 297 AsnMetLeuTrpSerIlePheAsnAspTyrArgArgTyrMetThrIleThrValLeuAsp 316
 DB 1507 GAGATACCAACAACCTATCAAAAT-----CTAACCAACCTAAGAT 1545
 QY 317 ThrIleSerGlnPheSerLeuTyrAspIleuArgTyrArgAspSerIleGlyIle 336
 DB 1546 ACTTTG-----TCCCTATGGGAT----- 1563
 QY 337 GluValIleGlyIleLeuAsnGluLeuThrArgGluIleTyrThrThrGluIleAsnPhe 356
 DB 1564 -----AATGAACTTGGGGCATTTACCCCAAAACTGCACGCTC 1605
 QY 357 AspArgLeuProGlnLeuArgValGlnProAsnLeuAlaThrMetGluTyr----- 373
 DB 1606 ACCAAGATGAAATACCTTAGCTTAGACAGCAACAACCTTAGCGAAATCCCTGGGTGT 1665
 QY 374 -----AsnLeuThrArgAlaSerPheLeuPheSerPheLeuGlu----- 387
 DB 1666 TTATCAATCTAACAATAATGAA-----AACTTACCTATACCAAAACCAAGTCACGGGT 1722
 QY 388 -----GlnPheIlePheTyrThrGluAsn 395
 DB 1723 TCCATCCCAAGAAATAGGAATGCTTCCCAACCTCAGTTTGTGAGCTAACCAACAC 1782
 QY 396 ThrAsnPheGlyAsnArgLeuValGlyIleSerAsnArgAspAlaProThrTyrSerAsn 415
 DB 1783 ACCTTAGCGGGAGATATCCACAGCTATCAAAAT-----CTAACCAAC 1827
 QY 416 ThrIleThrGluThrLeuTyrGlyIleArgThrGlySerProThrThrIleArg 435
 DB 1828 CTAGCATATTTGCTCATGGGTAAATGAACTTGGGGCTATACCCCA----- 1878
 QY 436 ProPheGluSerTyrIleValSerIleValThrAspArgIleSerProProValSerPro 455
 DB 1879 -----AACTCTGCATGCTCACCAAGATCCAAATCTTGAATCTTAGCAGT 1923
 QY 456 IleGlnProHisPheIleIleAsnGln-----IleGluLeuTyrIleAsnGlySer 472
 DB 1924 TCTCTTCTCGGAATTCGAAATCTCAGCGGCATGACAGACCTTGGCTGATTAATAC 1983
 QY 473 Ser-----AsnAsnThrLeuTyrTyrSerAlaGlyIleSerLeuSerAsnTyrGln 489
 DB 1984 TCATTTTCAGACATTTGCTGCAGAAATGATGATGGGTGGCAGACTCAAAACTTTCAAG 2043
 QY 490 -----AsnThrThrPhePheGlnPheProArgIleLysAspCys----- 502
 DB 2044 ATTGTGGGAATGCTTCGATGAGCCCATTCACAGAGAGCTTGAAGACATGTACAGCTTG 2103
 QY 503 -----AsnLeuValIleAspProGlyCysSerProAsnPheAsnAsn 516
 DB 2104 GTTAAACTTCCGTGTACATTAACCTGCTTACCGGAGACATATCTGAACATTTTGAAGTG 2163
 QY 517 TyrSerHisIleLeuSer-----HisPheSerLeuPheThrTyr 529
 DB 2164 TATCCCATCTCAAAATCGGTGAGCTTGTCAATACATAGTCTTGGGAGATCTGACCA 2223
 QY 530 SerTyrValIleGlyLeuLeuGlnIleLeuAsp-----ThrGly 543
 DB 2224 AATGGGTGCTGATGCTCCAGCTTGAAGAAATGATTTTCACAAAACATGATCACAGGT 2283
 QY 544 ValLeuGlyTyrPheHisSerValAspArgTyrAsnAlaIleSerAspIleIle 563
 DB 2284 TTACTGAGACTTATCAT-----AATTAATATCAGTGTGAG----- 2319
 QY 564 ThrMetIleProAlaIleIleGlyAsn-----AsnLeuAspThr 576
 DB 2320 -----ATTCCAGCAGATTTGGAATTCGAAAGCTTATATAGATAAATTATATCATTC 2373

QY 577 AsnSerIleValIleGluGlyProGlyHisThrGlyIleAsnLeuValTyrLeuGlnSer 596
 DB 2374 AAT-----CAGCTTCAGAGATTTTATCCAGCAGCAGCTAGGAAGTTAAGCAAT 2421
 QY 597 GlnGlyArgLeuGluIleThrCysGluThrProAsnSerThrGlnSerTyrPheIleArg 616
 DB 2422 TTAGGGTACCTTGAATGTATCAAGAAC----- 2448
 QY 617 LeuArgTyrAlaThrAsnGlyAlaGlyAsnThrLeuProAsnIleSerLeuThrIlePro 636
 DB 2449 -----AATTTAGTGCACCAATACCC 2469
 QY 637 GluValIleGly-----IleProProGln 644
 DB 2470 GATGAACATGAGGAGCATGCATGCATTCAGAGAACTTCCATCATTCGTAAGCGCC 2529
 QY 645 ArgLeuAsnAsnThrPheSerGlyThrAsnTyrAsnAsnLeuGlnTyrGlyAspPheGly 664
 DB 2530 ATGTTGAGCTTAAACAGTTCTTGATGTGCTCTCAACCAACTTGGA----- 2574
 QY 665 TyrPheGlnPheProSerThrValThrLeuProLeuAsnArgAsnIlePro---PheIle 683
 DB 2575 ---GGACTCTCCAGCAGACATCTACTTCAA---AATGCATCAATTAAGCTGTATTAT 2628
 QY 684 PheAsnArgAlaAspValSerAsnSer----- 692
 DB 2629 CACATTAAGCTTATAGCATCTGTGCATCAATTAAGCCAAAGCTACTAAGTTGCTACGCCA 2688
 QY 693 -----IleLeuIleIleAspIleGlyIlePheIleProIleThrSerSerMet 708
 DB 2689 ATTTGTTCTGTGTGTGCTATTCATTTCTGGGTCAAAATATTATGTATCAAAAGCTTGT 2748
 QY 709 HisGlnAsnArgGluIleGlyIleGluIleGlnThrIleAsnThrPhePhe 728
 DB 2749 CATTAACAGAGAAACCAACAGCAAGCTCTAGTCTATTACTGTTCAGAGAACATGTT 2808
 QY 729 Thr 729
 DB 2809 TCT 2811
 RESULT 4
 AG392993/c 1754 bp DNA linear GSS 21-DEC-2004
 LOCUS Mus musculus molossinus DNA, clone:MSMg01-211C07.T0, genomic survey
 DEFINITION
 ACCESSION AG392993 1754 bp DNA linear GSS 21-DEC-2004
 VERSION AG392993
 KEYWORDS
 SOURCE Mus musculus molossinus (Japanese wild mouse)
 ORGANISM Mus musculus molossinus
 Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
 Sciurognathi; Muridae; Murinae; Mus.
 REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 PUBMED
 REFERENCE
 AUTHORS
 DIRECT SUBMISSION
 JOURNAL
 15574823
 2 (bases 1 to 1754)
 Hattori,M., Toyoda,A., Noguchi,H., Kojima,T. and Sakaki,Y.
 Contribution of Asian mouse subspecies Mus musculus molossinus to
 genomic constitution of strain C57BL/6J, as defined by BAC-end
 sequence-SNP analysis
 Genome Res. 14 (12), 2439-2447 (2004)
 Submitted (17-NOV-2003) Maastricht Hattori, The Institute of Physical
 and Chemical Research (RIKEN), Genomic Sciences Center (GSC),
 1-7-22 Suenro-chou,Tsukumi-Ku, Yokohama, Kanagawa, 230-0045, Japan
 (E-mail:hattori@gsc.riken.jp, URL:http://hgp.gsc.riken.go.jp/
 Tel:81-45-503-9111, Fax:81-45-503-9170)
 Clones are derived from the mouse BAC library MSMg01. For BAC
 library availability, please contact Kuniya Abe (abe@rtc.riken.jp).

Tsukuba Institute, Bio Resource Center,
The Institute of Physical and Chemical Research (RIKEN) 3-1-1
Koyada, Tsukuba, 305-0074 Japan
phone: 81-298-36-9189, fax: 81-298-36-9199
e-mail: abee@rc.riken.jp

```

/tissue_type="mixture of kidney and spleen"
/clone_1ib="MSMG01 Mouse Male BAC Library"

```

US-10-783-417-2 (1-735) X AG392993 (1-1754)

QY	413	YyseAenThrlleThgGluThleuYr-----GluGluXgThGlySerProHr	430
Db	1748	TACACATGACATACATACATGAAACGATCTACTTCAACAAACCGTACCAATGCATCATCAATTC	1685
QY	431	ThrlYsThrlle-ArgProPheGluSerTyrrYsValserIleValThraspArgLase	450
Db	1688	ACA---ACGATANTTCCTCTTAACACACATCATATTCACAACTCAACCTTCATGAA--	1633
QY	450	rProProValserProIleGlnProHisPheIleIleangInIleGluLeu----	468
Db	1633	---CCACATTAACACATACACATCACTCTTACTCTTAAACACACAGACTTCATCATTA	1578
QY	468	rLeuansGlySerSerAsnaAenThrlleuYsrySerhLaglyGlySerLeuSerAsnTy	488
Db	1577	TACACACATTCATCATCTCTCTTACA-----TCATCTCACAGTA	1533
QY	488	rGlnAsn-----ThThrPhePheGln-PheProArglySlyAspC	502
Db	1538	TCACATGACACGACAGACACATCATCTGCTCGTCAAGATTAATCGCATCAACAGAACTG	1477
QY	502	yAsnIleuVal-----IleAspProGlyCySerProAsnPheAsnAntySer----	518
Db	1478	CCTCTCTATCATCTACTACTACAGTACATGCTGCACCAACGTAACATATCATCATGAA	1418
QY	518	-----HisIleLeuSerHisPheSerLeuPheThrTySerT	531
Db	1418	ACACACACACACTCAACTACATACATATTCCTCATACATCACTCACTC-----	1377
QY	531	yrValIleGlyLeuGlnIleuGlnIleLeuAspThGlyValleuGlyTrpThrHisSerS	551
Db	1371	-----ACTCATCTAC	1367
QY	551	erValAspArgTyrr-AsnAlaIleSerAspIlyIleIleThrMetIleProAlaIleYs	570
Db	1361	ANTACAAATATCATATCATGCTGCTACTCTCTGTGACACATACAAACATCTCAACTCAACAA	1302
QY	571	GlyAsnAsnIleAspThrAsnSerlyValIleGluGlyProGlyHisIleThrclyGlyAsn	590
Db	1301	TCAGTATACACAAACACATATTCACATCTCATTAACATACCAACCAACATGAGGCATCA	1247

Oy		LeuVal ¹ ThyrLeugInser-----GlnIly	598
Dd	1241	CTCAATTCATTAACAACATCATACTACTTCAATAATCAACATCTTAACAACCGTGCGGGGT	1187
Oy	599	ArgLeuglu-----IlEthrCysgluthrProamSerThrGlInser	612
Dd	1181	CGAATCTCTCATGCACACAAACATCAACAACAACA- CAGAGCCCCCACTCCAAAAGCTCCA	1122
Oy	613	ThyrHeillea ¹ gLeuaargTyrlalEthrasnIylaglYasnthlrProasnIIeser	632
Dd	1122	TAT-----CTACATATTAACAACAACATCTATCTTAATTCATCTTTAAAGTAAC	1077
Oy	633	LeuthrlleProglIvalIIleglylIeProProglInargLeuaAsnaenthPhsesexily	652
Dd	1071	--ACATGTCGA-----ACCACATCCCATATTAAGAACAATATTAGCCCC	1033
Oy	653	ThrasnTYzraanaenleugIntyrglyla ¹ a ¹ rPhegllytgrRheglInpheProserThryal	672
Dd	1029	ACATTCCT--NATCAA-----CACACTCATATTCCAAACCAATAG	993
Oy	673	ThrlaupProleuaanargAnllEProPheillePheena ¹ rglalaaPvalBerxanser	692
Dd	992	AACACTCCAAATCATTAACAAATCCCAACATGACACAAACCTTAATTAATCTATATCAG	933
Oy	693	IleleuileilaaplysIIegluPheilleProIIethrSerSermethIcglIn-----	710
Dd	932	ATATTTCATACCTCCACATGCCAACAAACATCAACACATCAACAATACCGCTACCAATAAC	873
Oy	711	Asna ¹ rgluysgslInlyleugInThrlleglInThrys	723
Dd	872	AACAACGAACAACAACATCAATCAAAACCAATTAATCAANAAG	834

RESULT 5
CNS0A8VH

LOCUS CNG0A8VH 2032 bp mRNA linear HTC 06-FEB-2004
DEFINITION Arabidopsis thaliana Full-length cDNA complete sequence from clone
GSI:FPB40Z03 of Flowers and buds of strain col-0 of Arabidopsis
thaliana (thale cress).
ACCESSION BX819057
VERSION BX819057.1 GI:42467943

KEYWORDS	HTC; GS/LT cDNA.
SOURCE	Arabidopsis thaliana (thale cress)
ORGANISM	Arabidopsis thaliana
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophytes; Magnoliophyta; eudicotyledons; core eudicotyledons; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi.
REFERENCE	1 (bases 1 to 2032)
AUTHORS	Castelli, V., Aury, J.M., Jillion, O., Winkler, P., Clepet, C., Menard, M., Guaud, C., Quetier, F., Scarpelli, C., Schachter, V., Temple, G., Caboche, M., Weissendbach, J. and Salamouch, M.
TITLE	Whole Genome Sequence Comparisons and Full-Length cDNA Sequences

COMMENT

Submitted (18-NOV-2003) Genoscope - Centre National de Séquençage BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr)

The sequences are based on single pass reads. Life Technologies (a division of Invitrogen) members carried out full-length libraries construction ; Temple G. Genoscope members carried out sequencing and annotation ; Castelli V., Aury J.M., Jaillon O., Wincker P., Menard M., Craud C., Schachter V., Weissenbach J., Salamoubat M. URGV INRA ; Clepet C., Caboche M.

Annotation is based on the June 2003 version of the Arabidopsis Genome released by MIPS (Munich Information center for Protein Sequences) . 5 prime and 3 prime are assembled with Phrap. http://www.genoscope.cns.fr/externe/sequences/Banque_Projet_EF/Fulllength

<http://www.genoscope.cns.fr/cgi-bin/ggb/ggb?source=Arabidopsis>.

FEATURES
source
Location/Qualifiers
1..2032
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/db_xref="taxon:3702"
/clone="GSLTFB402E03"
/tissue_type="flowers and buds"
/ecotype="Col-0"
/plasmid="pCMVSPORT_6"
1..2032
/gene="At2g35230"
/gene="At2g35230"

ORIGIN

Alignment Scores:
Pred. No.: 0.0582 Length: 2032
Score: 121.00 Matches: 95
Percent Similarity: 35.2% Conservative: 72
Best Local Similarity: 20.0% Mismatches: 168
Query Match: 3.1% Indels: 139
Gaps: 23

US-10-783-417-2 (1-735) x CNS0A8VN (1-2032)

QY 306 TyrAgaGgTyrMeethrIleThrValIleuAspThrIleSerGlnPheSerLeuTyrAap 325
|||
DB 480 TACTATAGGATTTGATTTTGTGAGATGATAGG----- 515
QY 326 IleLysArgTyrAgaSperIleGly-----GlyIleGluValIleGlyIle 341
|||
DB 516 ---CCTAGACAAATGATCATTTGGGTGATAGATAGGAAAAATATCAGA----- 566
QY 342 LysAenGluLeuThrArgGluIleTyrThrThrGluIleAsnPheAspArgLeuProGln 361
|||
DB 567 AAGAGCTCTTCATCAATGACATTCCTGCTGCAAGTACTAGCAACCGTCCCTCCAG 626
QY 362 LeuArgValIleProAsnLeuAlaThrMetGluTyrAsnLeuThrArgIleSerPheLys 381
|||
DB 627 CTTCACACACAGCTCAGGTT-----TATACATTAAGCAAGAACGACTTTTAA 674
QY 382 LeuPheSerPheLeuGlnPheIlePheTyrThrGluAsnThrAsnDheGlyAsnArg 401
|||
DB 675 ---AGTATGTTTACAGCAG----- 689
QY 402 LeuValGlyIleSerAsnArgAla-----ProThrTyrSerAsnThrIleThr 418
|||
DB 690 CTAAAGGGGTTCTCAACAGGTGAAGCTTCTCGGCTCTCAAGAACCTTCGCTA--- 746
QY 419 GluThrLeuTyrGlyLeuArgThrGlySerProThrThrIleArgProPheGlu 438
|||
DB 747 -----AGGCTTCAAGATACACGGTTGACGGGATCAGACCA----- 782
QY 439 SerTyrLysValSerIleValThr-----AspArgGlnSerProProValSerProIle 456
|||
DB 783 -----TCCCCCTTAAGCATTAACCCGGCTCGGTTCTCTCCCTTCAG 830
QY 457 GlnProHisPheIleIleAsnGlnIle-----GluLeuTyrLeuAsnGly 471
|||
DB 831 GCTCCGACAGTCTCATCTCAGTTGCTAGACAGCTCCACACACGACCTTCCAC 890
QY 472 SerSerAsnAsnThrLeuTyrTyrSerAlaGlyIleSerLeuSerAsnTyrGlnAsnThr 491
|||
DB 891 AAACACACACACACATGATGGGTACACAGGACCAAGTTTGGTGAATACAG----- 944
QY 492 ThrPhePheGlnPheProArgLysLysAspCys-----AsnLeuValIle 506
|||
DB 945 -----CTGAGTTTCCCGTCTCAGAGATATAGGTTATCTCAAGTTAACTTGGAGATT 998
QY 507 AspProGlyCysSerProAsnPheAsnAsnTyrSerHisIleLeuSerHisPheSerLeu 526
|||
DB 999 CAG-----GACCAATGCTAACCAATGACGCGAGTCAATGACGAC----- 1040
QY 527 PheThrTyrSerTyrValIleGlyLeuGlnLeuGlnIleLeuAspThrGlyValLeuGly 546

DB 1041 -----GGCTTATATCCAGAGTATGACAGGACCATATG--- 1076
QY 547 TrpThrIleSerSerValAspArgTyrAsnAlaIleSerAspLysIleIleThrMetIle 566
DB 1076 ----- 1076
QY 567 ProAlaIleLysGlyAsnAsn-----LeuAspThrAsnSerLysValIleGlu 582
|||
DB 1077 -----TGCACAGTATATGACAGACCATATATGCCAGTATATGACAGGCGCATAT 1130
QY 583 GlyProGlyHisIleThrGlyLysLeuValTyrLeuGlnSerGlnGlyArgLeuGluIle 602
|||
DB 1131 ATCCAGGTCATGACAGGCGTATATATGACAGTATATGACAGTATATGACAGTATAT 1190
QY 603 ThrCysGluThrProAsnSerThrGlnSerTyrPheIleArgLeuArgTyrAlaThrAsn 622
DB 1191 CAACCTCAA---CCTCAACCTCAGACATATATGATGCTGACCGGACCTGATATGAC 1247
QY 623 GlyAlaGlyAsnThrLeuProAsnIleSerLeuThrIleProGlyValIle----- 639
DB 1248 ATGACAGGCGCTTCAACCTCAACATATGATCTACACACCGGCTATGCTTACGCCA 1307
QY 640 -----GlyIleProGlnArgLeuAsnAsnThrPheSerGlyThr----- 653
DB 1308 GTGCTCATATCTTCTCTCTCGGTTCAATGCTCCGTACCTGTAACGCTACTAG 1367
QY 654 -----AsnTyrAsnAsnLeuGlnTyrGlyAspPhe 663
DB 1368 CCTCTCCCATGTTCACTGACATGATATGATGATGATTTCTTCTTCATATATACGCTTT 1427
QY 664 GlyTyrPheGlnPheProSerThrValThrIle---ProLeuAsnArgAsnIleProPhe 682
DB 1428 GACACATGACATCACTATATCCCATTTCTTCAACATCTTCAACCGGTTACCCGAAT 1487
QY 683 IlePheAsn-ArgAlaAspValSerAsn-----SerIleLeuIleI 696
|||
DB 1488 ATGTTCTCTCCAGATACCGCTTATACCATGCTGTCACAGAGTTCAAGTATCTTC- 1542
QY 696 eAspLysIleGluPheIleProIleThrSerSerMetHisGlnAsnArgGluGlyIle 716
|||
DB 1543 -----AACCACTTACCCCAAACTTCTATTTCAAAATGCTCAACAGAGAGT 1592
QY 716 eLeuGluThrIleGlnIleThrIleValIleAsnThrPhePheThr 729
|||
DB 1593 CTGGAACCGGTGACAGTCAAGTCAAGGTCTCTCTCAAC 1632

RESULT 6
CL967047
LOCUS 6777 bp DNA linear GSS 21-SEP-2004
DEFINITION Oryza sativa (indica cultivar-group) genomic, genomic survey sequence.
ACCESSION CL967047
VERSION CL967047.1 GI:52388742
KEYWORDS GSS.
SOURCE Oryza sativa (indica cultivar-group)
ORGANISM Oryza sativa (indica cultivar-group)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzoideae; Oryza.
REFERENCE 1 (bases 1 to 6777)
Ma, L., Wang, J., Chen, C., Liu, X., Su, N., Li, L., Wang, X., Cao, M.,
Jiao, Y., Sun, N., Zhang, X., Bao, J., Sun, D., Zhao, H., Yuan, L.,
Wong, G.K.S., Deng, X.W. and Wang, J.
An analysis of transcriptional regulation of the rice genome and
its comparison to Arabidopsis
Unpublished (2004)
JOURNAL
COMMENT
Contact: Chen Chen
Department of Bioinformatic
Beijing Institute of Genomics
Chinese Academy of Sciences, Beijing 101300, China
Tel: 86-10-80481559
Fax: 86-10-80486676
Email: chenchen@genomics.org.cn

QY 548 hTyrSerTyrValIleGlyLeuGlnIleLeuAspThrGly----- 543
 Db 1238 CTGAGGCAATCAAGAAATGCTCCAGCTAGACTCTGATTTGGCATTTACTGCATTA 1297
 QY 544 -----ValleuGlyTPT 548
 Db 1298 AGGTGAAGATGACAGAGTCCCTTCAGAACCTTCATTTTGAAGGTGGAATCTCT 1357
 QY 548 hTieSerSerValAspArgTyrAsnAlaIleSerAspIleIleThrMetIleProA 568
 Db 1358 CCCACAGCTCCTTATC-----ATCAGACGTGAGAGCTCTTCATAGGCTGCCAG 1408
 QY 568 laIle-----LysGlyAsnAsnLeuAspThrA 577
 Db 1409 CACTCCAGCATTTGAACCTTACAGAGGAATCATTTCACAAAGGAATATCCAAAGACCA 1468
 QY 577 hTieSerValIleGlyLeuGlyProGlyHieThrGlyGlyAsnLeuValTyrLeuGlnSerG 597
 Db 1469 ACTCA-----CTTCAGACAC 1483
 QY 597 InGlyArgLeuGlyIleThr-----CysGlyThrProAsnSerThrGlnSerT 613
 Db 1484 TGGGAAGACTAAGAAATCTGTTTATCTTTTGATCTCTTCATTCATGACCGACAG 1543
 QY 613 YPheIleArgLeuArgTyrAlaThr----- 621
 Db 1544 CCTTCACCGCTGTAAGATGATGATATCATGTAGACTGACGACAGGCTGACATCCA 1603
 QY 622 -----AsnGlyAlaGlyAsnThrL 628
 Db 1604 GTAGCATTTGAGGCTTACTACTCTTAAAGGGATCTTACTCTCAATCTGGCTCAAT--- 1659
 QY 628 eUProAsnIleSerLeuThrIleProGlyValIleGlyIleProProGlnArg 645
 Db 1660 -----CACATCAGCATCTCTTACCCAGCTCTCTCCCATCTTGTCCACGACG 1707

RESULT 8
 AK089255
 LOCUS
 DEFINITION
 AK089255 2157 bp mRNA linear HTC 03-APR-2004
 Mus musculus NOD-derived CD1c +ve dendritic cells cDNA, RIKEN
 full-length enriched library, clone: F630107B15 product: lymphocyte
 antigen 78, full insert sequence.
 AK089255
 VERSION
 AK089255.1 GI:26354375
 HTCC, CAP trapper.
 SOURCE
 Mus musculus (house mouse)
 ORGANISM
 Mus musculus
 Buxerjota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
 Sciurognathi; Muridae; Murinae; Mus.

REFERENCE
 AUTHORS
 1 Carninci, P. and Hayashizaki, Y.
 TITLE
 High-efficiency full-length cDNA cloning
 JOURNAL
 Meth. Enzymol. 303, 19-44 (1999)
 PUBMED
 10349636
 2
 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
 Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
 Normalization and subtraction of cap-trapper-selected cDNAs to
 prepare full-length cDNA libraries for rapid discovery of new genes
 Genome Res. 10 (10), 1617-1630 (2000)

JOURNAL
 PUBMED
 11042159
 3
 Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,
 Konno, H., Akiyama, J., Nishi, K., Katsunai, T., Tashiro, H., Itoh, M.,
 Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,
 Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,
 Fujiwara, S., Inoue, K., Togawa, Y., Izawa, K., Ohara, S., Watabiki, M.,
 Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J.,
 Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
 RIKEN integrated sequence analysis (RISA) system-384-format
 sequencing pipeline with 384 multicapillary sequencer

JOURNAL
 PUBMED
 11076861
 4
 The RIKEN Genome Exploration Research Group Phase II Team and the
 FANTOM Consortium.
 TITLE
 Functional annotation of a full-length mouse cDNA collection
 JOURNAL
 Nature 409, 685-690 (2001)
 REFERENCE
 AUTHORS
 5
 The FANTOM Consortium and the RIKEN Genome Exploration Research
 Group Phase I & II Team.
 TITLE
 Analysis of the mouse transcriptome based on functional annotation
 of 60,770 full-length cDNAs
 JOURNAL
 Nature 420, 563-573 (2002)
 REFERENCE
 AUTHORS
 6 (bases 1 to 2157)
 Adachi, J., Aizawa, K., Akimura, T., Arikawa, T., Bono, H., Carninci, P.,
 Fukuda, S., Furuta, M., Hanagaki, T., Hara, A., Hashizume, M.,
 Hayashida, K., Hayatsu, N., Hiramoto, K., Hirooka, T., Hirozane, T.,
 Hori, P., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kanakawa, T.,
 Katoh, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M.,
 Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M.,
 Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N.,
 Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N.,
 Sano, H., Sasaki, D., Shibata, K., Shingawa, A., Shiraki, T.,
 Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, P., Takaku-Akahira, S.,
 Takeda, Y., Tanaka, T., Tomaru, A., Toyota, T., Yasunishi, A.,
 Muramatsu, M. and Hayashizaki, Y.
 TITLE
 Direct Submission
 JOURNAL
 Submitted (16-APR-2002) Yoshihide Hayashizaki, The Institute of
 Physical and Chemical Research (RIKEN), Laboratory for Genome
 Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
 RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
 Kanagawa 230-0045, Japan (E-mail: genome-ree@gscc.riken.jp,
 URL: http://genome.gsc.riken.jp/, Tel: 81-45-503-9222,
 Fax: 81-45-503-9216)

COMMENT
 cDNA library was prepared and sequenced in Mouse Genome
 Encyclopedia Project of Genome Exploration Research Group in Riken
 Genomic Sciences Center and Genome Science Laboratory in Riken.
 Division of Experimental Animal Research in Riken contributed to
 prepare mouse tissues.
 Tissues were provided by Dr. John Todd (Dept. of Medical Genetics
 Wellcome Trust Centre for Molecular Mechanisms in Disease Wellcome
 Trust/MRC building Addenbrookes Hospital Cambridge) whose
 assistance we gratefully acknowledge.
 Please visit our web site for further details.
 URL: http://genome.gsc.riken.jp/
 URL: http://fantom.gsc.riken.jp/.

FEATURES
 source
 location/Qualifiers
 1..2157
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="NOD"
 /db_xref="PANTOM:DB:F630107B15"
 /db_xref="taxon:10090"
 /clone="F630107B15"
 /cell_type="NOD-derived CD1c +ve dendritic cells"
 /clone_11b="RIKEN full-length enriched mouse cDNA library"
 134..2119
 /note="unnamed protein product; lymphocyte antigen 78
 (MG) MG1:1194924, GI:NM_008533, evidence: BLASTX, 99%,
 match=2136)
 putative"
 /codon_start=1
 /protein_id="BAC40816.1"
 /db_xref="GI:26354376"
 /translation="MAADPISCFVVALPASCRAATSSDQCIKRYNKYTCNGL
 NEITGLPNSTECHAFSPNVLPFTONTTFSLLINTFIDLRQCIYTHETPQGR
 LDTLVLTNPLIPAEATLSPKALKHLFTQTGISIDTLPINOKTLBSLYASNH
 ISSIKLPKGPPTKXKLVDFQNNAIHYISKDMSIQATVLSLNKNDIAGETLGA
 FDSAVFGSINFGQNTLVIFKGLKNSITGSLMTGTFEDMDDEIDSPVFGSLCMSV
 BSINOKHYFNISNTNFCPSGLOEIDLTLTHSRLSPGVGSTLKVLVSAKFV
 NLCOISANFSLTHLSIKNTKRLLELGTCLNENLRRLDLSHDIETSDCCVLT
 RLNSHLSLNTSVPEPLSKTEARFECQLBLDLAFLRYKDAQSPQNLHLKVL
 NLHSLDLISSEQLFDGLPALQHLNLQNHPPKQNIQKTNLSLQTLGRLELTVLSPCDL

ORIGIN

SSIDHAFTSLKMMNHVDLSHNLNTSSIEALSHLKGIYLAIVASNISIIPLSLPIL
SOORLIRKONPLDCTCSNIYFLEMYKEMOKLEDEPDLCENPPLLRGRILSDVTL
CSMAAVGIEFLIYFLLVPLLIIFAVKYFLRMKYKH1"

Alignment Scores:

Pred. No.:	0.188	Length:	2157
Score:	117.00	Matches:	114
Percent Similarity:	31.7%	Conservative:	76
Best Local Similarity:	19.0%	Mismatches:	190
Query Match:	3.0%	Indels:	221
DB:	4	Gaps:	22

US-10-783-417-2 (1-735) x AK089255 (1-2157)

```
QY 201 AenValHiasnAspRheleArgGluileProGly----- 212
DB 245 AACTGTGAAATTTAGGTCTCAATGAAATTCCTGGACCTTTACCAAACTCCAGAGATGT 304
QY 213 -----PheGlnleuGlnThrTyLysThrleuLeuProIleTyrlaGlnlaala 230
DB 305 TTGGATTCAGCTTTATATGCTTCCCTACCATTCAAAACAGACCTTCAGACATTTATC 364
QY 231 AenPheHiasnleuLeuGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 246
DB 365 AATCT-CACTCTCTCGATTTACAGAGTCCAGATTTACTGATACATGAAGATCTTT 423
QY 246 ----- 246
DB 424 CCMAACCAACATCGGTAGACACACTGTACTACCGCAATCCCTGATATTATTCG 483
QY 247 -----ThrAsnAlaAspIleHiasProSerGlnIleGluProAsnAlaGly 261
DB 484 AAGAGACAGACTTATAGTGGCTTAAGCCTTAAGCATTCGTCTTCATCCAAACAGAGAT 543
QY 262 -ThrsSerAspTyLysLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 281
DB 544 ATCCAGATATGATTTATTCCTCCACTGCACATCAAAAACCTTGAA--AGTCTATCT 600
QY 281 baIaAsnThrTyLysArgThrGlyLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 301
DB 601 CGGACCAACCAATATTCCTCCACTTAAGCTCCCAAGGTTTCCCAACAGAGACCTGA 660
QY 301 rIlePhe-----AsnAspTyLysArgThrGlyMetThrIleThrValLeuAspThrI 318
DB 661 GGCTCGATTTTCAGAAATATATCTATTCATTCCTGCTAAAGAAATATGAGCTCTCT 720
QY 318 eSerGlnPheSerLeuTyLysAspIleLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 338
DB 721 CCAGCAGACCACTAATCTGAGCTTAACTTAATGAATGAATGACATTCGACGAATGAGCT 780
QY 338 1----- 338
DB 781 GGGGGCTTTCAGACTGCTCTTCCAAAGTTTAACTTTGAGGAGACTCAGAACTTGCT 840
QY 339 -----LysGlyIleLysAsnGlnLeuThrArgGluIleTyThrThrGluIleAs 355
DB 841 AGTTATCTTCAAGGGTTTGAAGAACTTAAGATCCAGTCTCTTGGCTGGGACATTTGA 900
QY 355 n-----PheAspArgLeuProGlnleuArg 364
DB 901 GGACATGATGACGAAGATATTAATCTCCCTGCTTTGAGGGCTCTGGAATATGCTGT 960
QY 364 IglNProAsnLeuAlaThrMetGluTyL-----AsnLeuThrArgAlaSerPheLeu 382
DB 961 GGGAGCATCAACCTTACAGAGCATTAATTCCTCAACATTCCTCCCAACATTCATTCG 1020
QY 382 upPheSerPheLeuGlnPhe-----IlePheTyThrGlnAsnThrAsnPheLeuAsn 401
DB 1021 CTTCACTGGCTTCAGAGAACTGACCTTAACAGCCTCACTGAGTGAATTCCTTCCTG 1080
QY 401 gIeuValGlyIleSerAsnArgAspAlaProThrTyLysAsnThrIleThrGluThrIle 421
```

```
DB 1081 ACTTGGGACCTAAC----- 1096
QY 421 uTyGlyGluArgThrGlySerProThrThrLysThrIleArgProPheGlu-SerTyL 441
DB 1097 -----ACACTTAAGAAATTAAGTCTCAGTGC 1122
QY 441 yValSerIle---ValThrAspArgGlnSerProValSerProIleGlnProHiasp 460
DB 1123 AATAAGTTTGAAGATTTTGGCAATATCAGTGTCCATTTCCCTCC-CTTACTACAC 1181
QY 460 heIleIleAsnGlnIleGluLeuTyLysAsnGlySerSerAsnAsnThrLeuLysTyS 480
DB 1182 TTTCCATC-----AAGGCGACAGAGAGACTTG 1211
QY 480 eAlaGly---GlySerLeuSerAsnTyLysAsnThrThrPheGlnPhePro---- 497
DB 1212 AGCTGGTCTGAGCTGCTTTGAAACCTTAAGAAATTCGCGCAACTTGAACCTCAGC 1271
QY 498 -----ArgLysLysAspCysAsnLeuValIleAspProGlyCysSerProAsnPhe 515
DB 1272 ATGACATTAAGAACTTCTGATGT-----TCCAACTCCCACTCC 1310
QY 515 snAsnTySerHisIleLeuSer-----HisPheSerLeuPheT 528
DB 1311 GAAACCTGTCTACCTTAACAGACCTGAACCTGATACATGAACCCCTGAGCTCAAG 1370
QY 528 hTyLysSerTyValIleGlyLeuGlnleuGlnIleLeuAspThrGly----- 543
DB 1371 CGAGGCAATCAAGAAATCCCTCAGCTGAACTCTTAATTTGGCATTTACTCGACTAA 1430
QY 544 -----ValLeuGlyTrpT 548
DB 1431 AGGTAAAGATGACAGAGTCCCTTCCAGAACTCATCTTTGAAGTGTCTCATCTCT 1490
QY 548 hHisSerSerValAspArgTyLysAlaIleSerAspLysIleIleThrMetIlePro 568
DB 1491 CCCAGAGCTCTCTGAC-----ATCAGAGTGAAGAGCTCTTCGATGGCTCCAG 1541
QY 568 laIle-----LysGlyAsnAsnLeuAspThr 577
DB 1542 CACTCAGCATTTGAACCTTAACAGGAATCACTTCCAAAGGAATATTCAAAGACCA 1601
QY 577 snSerLysValIleGluGlyProGlyHisThrGlyLysLeuValTyLysGlnSerG 597
DB 1602 ACTCA-----CTTCAAGAC 1616
QY 597 IngIlyArgLeuGluIleThr-----CysGluThrProAsnSerThrGlnSerT 613
DB 1617 TGGGAAGACTTAAGAAATCCGTTTATATCTTTTGATGTCTCTCCATTTGACAGAG 1676
QY 613 yPheIleArgLeuArgTyLAlaThr----- 621
DB 1677 CCTTACCAAGCTGAAGATGAATCATGTAGACCTGAGTCAACAGAGCTGACATCCA 1736
QY 622 -----AsnGlyAlaGlyAsnThrL 628
DB 1737 GTAGCATTAAGGCTTTAGTATCTTAAGGGATCTCACTCAATCTGGGCTCCAAAT- 1792
QY 628 euProAsnIleSerLeuThrIleProGlyValIleGlyIleProProGlnArg 645
DB 1793 -----CGATCAAGCATCATCATCCAGTCTCTCCATCTTGTGCCAGAG 1840
```

RESULT 9
AJ538005 1178 bp mRNA linear EST 01-JUL-2004
AJ538005 1A1 Timarcha balearica cDNA clone Timarcha6B7, mRNA
DEFINITION
sequence.
ACCESSION AJ538005
VERSION AJ538005.1 GI:40311007
KEYWORDS
SOURCE Timarcha balearica
ORGANISM Timarcha balearica
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;

Neoptera; Endopterygota; Coleoptera; Polyphaga; Cucujiformia;
Chrysomeloidea; Chrysomelidae; Chrysomelinae; Timarchini; Timarcha.
1 (bases 1 to 1178)
Gomez-Zurita, J., Kojilku, F., Theodorides, K. and Vogler, A. P.
Resources for a phylogenomic approach in leaf beetle (Coleoptera)
systematics
(in) Jolivet, P., Santiago-Blay, J. A. and Schmitt, M. (Eds.), NEW
DEVELOPMENTS IN THE BIOLOGY OF CHRYSOMELIDAE, 19-35 (2004)
Contact: Gomez-Zurita, J
Department of Entomology
The Natural History Museum
Cromwell Road, London SW7 5BD, United Kingdom.

COMMENT

FEATURES

source
location/Qualifiers
1..1178
/organism="Timarcha balearica"
/mol_type="rRNA"
/db_xref="taxon:79517"
/clone="Timarcha687"
/dev_stage="adult"
/clone_1lb="1A1"
/note="country: Spain; Balearic Islands, Mallorca, Balearics"

ORIGIN

Alignment Scores:
Pred. No.: 0.12 Length: 1178
Score: 115.00 Matches: 75
Percent Similarity: 34.9% Conservative: 35
Best Local Similarity: 23.8% Mismatches: 106
Query Match: 3.0% Indels: 99
DB: 1 Gaps: 15

US-10-783-417-2 (1-735) x AJ538005 (1-1178)

QY 470 AaNGlySeSerAaSnThLeuYsTYSerAlaGlyGly----- 483
DB 227 AATGGCTCAGCAACAC-----AATATATCTCTAATATATATATATCAAGTAATGCT 280
QY 484 -----SerLeuSerAaNTYrGlnAaNTThrPhePheGlnPhePro--Arg 498
DB 281 AATTATTAACCAAGTAACGGAACCTATCAAGAAACAATTGGACAGATTTCATCATG 340
QY 499 LysLYsAspCYaSnLeuValIleAspProGlyCYeSerProAsnPhaSnAsnTYSeR 518
DB 341 AATAACGATAGATATCAACGCTCAACGGAATCCCTTCCAAACATGATATATAACGG 400
QY 519 HisIleLeuSerHisPheSerLeuPheThrTYrSerTYrValIleGlyLeuGlnLeuGln 538
DB 400 ----- 400
QY 539 IleLeuAspThrGlyValLeuGly-----TrpThiHisSerSerValAspArgTYrAa 556
DB 401 -----GACCAAGGAATATATGGAATGCTATCCCATTCAGAAATAGTGATACAC 454
QY 557 AlaIleSerAspLYsIleIleThrMetIleProAlaIleLYsGlyAaSnLeuAsnLeuAspThr 576
DB 455 AATAATTAATCAACA-----AATAATCTGCCAGAAATATATGGAATTT 499
QY 577 AaNSerLYsValIleGlnGlyProGlyHisThrGlyGlyAaSnLeuValTYrLeuGlnSeR 596
DB 500 AATTCGACCAAGTAATGATATTCAGAAATAGTGATCAATCTTATCCATCACAGAAC 559
QY 597 Gln-----GlyArgLeuGlnIleThrCYsGlnThrProAsnSerThrGlnSerTYr 613
DB 560 AAAAATTATCCAGAAATATATGGAATAGCTCTCTCCATCCAGAAAT---AATACCTAT 616
QY 614 PheIleArgLeuArgTYrAlaThrAaGlyAlaGlyAaSnThrLeuProAsnIleSerLeu 633
DB 617 -----ACAGAAATAGTGGAATGCAATCTTATCATCACAAACAAAT 655
QY 634 ThrIleProGly-ValIleGlyIle--ProProGlnArgLeuAsnAsnThrPheSerGly 652
DB 656 AATTATCCAGAAATATATGGAATAGCTTCAACCTTCAACGAAATATATCTTATCCAGGA 715

QY 653 ThrAaNTYrAaSnLeuGlnTYrGlyAspPheGlyTYrPhe----- 666
DB 716 ACTAATGGAATATCTTACCAATCGAATAGTGTAATCTATCCAAAGATATATGTAT 775
QY 667 GlnPheProSerThrValThr----- 673
DB 776 AATTTCTCTTCAGTAATACCAAGTTCATATCCAGAACGATGGCAATATATCCAGGC 835
QY 674 LeuProLeuAaArgAaSnIleProPheIlePheAa-----ArgAlaAspValSer 690
DB 836 TTGAATATTAATTAAT 895
QY 691 AaNSerIleLeuIleIleAspLYsIle-----GlnPheIleProIleThr 705
DB 896 GGCAGCTGGGAAATATACGAACAGAAATGGGAACTTTCGGAAGTCAATCCCTAGCACT 955
QY 706 -----SerSerMetHisGlnAaSn----- 711
DB 956 CAACCAAAAATTTTACTGGCATACCGCTAGTGGGCCCACTTACATATAAATAATAGG 1015
QY 712 -----ArgGlnLYsGlnLYsLeuGlnThrIleGlnThr 722
DB 1016 ACGCCCAATGAAGAAGAGGTGTGAGTAATGTATAACG 1054

RESULT 10
CL960568 2322 bp DNA linear GSS 21-SEP-2004
LOCUS OaRCC004776 Oryza sativa Expressed Library Oryza sativa (indica
DERIVATION cultivar-group) genomic, genomic survey sequence.
ACCESSION CL960568
VERSION CL960568.1 GI:52375842
KEYWORDS
SOURCE
ORGANISM
Oryza sativa (indica cultivar-group)
Oryza sativa (indica cultivar-group)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzaceae; Oryza.
1 (bases 1 to 2322)
Ma, L., Wang, J., Chen, C., Liu, X., Su, N., Li, L., Wang, X., Cao, M.,
Jiao, Y., Sun, N., Zhang, X., Bao, J., Sun, D., Zhao, H., Yuan, L.,
Wong, G. K. S., Deng, X. W. and Wang, J.
An analysis of transcriptional regulation of the rice genome and
its comparison to Arabidopsis
Unpublished (2004)
Contact: Chen Chen
Department of Bioinformatic
Beijing Institute of Genomics
Chinese Academy of Sciences, Beijing 101300, China
Tel: 86-10-80481559
Fax: 86-10-80488676
Email: chenchen@genomics.org.cn
Rice genomic sequence.
Classes: exon-trapped.

FEATURES

source
location/Qualifiers
1..2322
/organism="Oryza sativa (indica cultivar-group)"
/mol_type="genomic DNA"
/db_xref="taxon:39946"
/clone_1lb="Oryza sativa Expressed Library"
/note="Oryza sativa exon trapped genomic sequences"

ORIGIN

Alignment Scores:
Pred. No.: 0.362 Length: 2322
Score: 115.00 Matches: 91
Percent Similarity: 35.7% Conservative: 74
Best Local Similarity: 19.7% Mismatches: 187
Query Match: 3.0% Indels: 110
DB: 10 Gaps: 19

US-10-783-417-2 (1-735) x CL960568 (1-2322)

OY	256	IIleGIuProAsnAlaAGLYThSerAspApyrTyLyvSleuAluLyGIuAsnIlePro	275
Db	628	TTAgATTTGGACAGACAGAAAGAAAGAACTGGCTGCTTCCGAGCAAGTACAGT	687
OY	276	LYSTySerAsnTyrcyAlaAsnThrTyArgThrGIyLeuLyvAsnLeuArgAspIu	295
Db	688	AGGATTTCTTAATTCGCTCAAGGATAGTTGGCAACAGATATGGATGCATTCGATAC	747
OY	296	ProAsnMetLySTyPserIle---PheAsnApyTyArgArgTyTMetThrIleThrVal	314
Db	748	AACTTGAGAGATTCACCGTTTCGAATTACAGAAATATCTTAAGATCATGCACCTTGTCGA	807
OY	315	LeuAspThrIleSerGIuPheSerLeuTyArgIle-----	326
Db	808	CTTCTCACTGATGAAGATTTGTGTTGCAGAAAATAGCTTTCTGTGTTCACAAATA	867
OY	326	-----	326
Db	868	ACTACGAAATCTGTACATATCAAGTGTCCCTCCCTACTAATGTGGGTGCTCATGAT	927
OY	327	LYsArgTyTArgAspSerIleGIyGIyIleGIuValLyvGIyIle-----LYsAsnGIu	344
Db	928	AAAAGGTTAGAAAGTTCATGGGCCCCAGATGATTAACGAATATCATCTTTACATAG	987
OY	345	LeuThrArgGIuIleTyThrThrGIuIleAsnPheAspArgLeuProGIuLeuArgVal	364
Db	988	CTGAGAGAGGCTGTAAAAGCTCATCATTTGGAAATGTGCACATGGTGAATCTTCAAA	1047
OY	365	GIuProAsn-----LeuAlaThrMetGIuTyrc	373
Db	1048	TGTCCAAATGCACATTAGACATGCACATTCGATCTGTGAAGTTGGCATTTCCAGAAATG	1107
OY	374	AsnLeuThrArgAlaSerPheLyvLeuPheSerPheLeuGIuGIuPheIlePheTyThr	393
Db	1108	AGGATATTACAGAGCTTCCTTTGTGTTGCATTTCAAGTTCGAAGCTGTATATCTGCTA	1167
OY	394	GIuAsnThrAsnPheGIuAsnArgLeu---ValGIyIleSerAsnArgAspAlaProThr	412
Db	1168	CTGACTTGCATGGATTAATGAATGATTTGCTATGTGATACGACGATTAAGCAATA	1227
OY	413	TyrcSerAsnThrIleThrGIuThrLeuTyrcGIuArgThnGIySerProThrThrLyv	432
Db	1228	ATGCATTACGATTATCTTCCAGATTTGAAGACGG-----GGGCGCTGAAGCCAAAG	1281
OY	433	ThrIleArgProPheGIuSerTyLyvAlaSerIleValThrAspArgGIuSerProPro	452
Db	1282	CTCACAGATATTTA-----GCACAGATTCGGTA	1311
OY	453	ValSerProIleGIuProIlePheIleIleAsnGIuIleGIuLeuTyTLeuAsnGIySer	472
Db	1312	TGTGATACCAATGAACTCATATATCTCAATATATCAAAATGATTTGGAAGGCGAGA	1371
OY	473	SerAsnAsnThr-----LeuLyvTyrcSerAlaGIyLyvSerLeuSerAsnTyrc	488
Db	1372	AATGATATTAATCTGATGTGACCTGTGAAGATATTCATTTGGAAGACCTCTTGTGTAT	1431
OY	489	GIuAsnThrThrPhePheGIuPheProArgLyvLyvAspCyAsnLeuValIleAspPro	508
Db	1432	ATTAACAT-----AAATATGCACCTATTGACAGATAT	1467
OY	509	GIyCySerProAsnPheAsnAsnTyrcSerHisIleLeuSerHisPheSerLeuPheThr	528
Db	1468	GGTATT-----GGTTATGCCGCTATGTCTCAGCATTAAGATTAATGTGAATATA	1515
OY	529	TyrcSerTyrcValIleGIyLeuGIuLeuGIuIleLeuAspThrArgValLeuGIyTyrcThr	548
Db	1516	ACAAACTTCTCATGTGTTAAGGAGAGCAACACTTACCAATGGCGCTCTGCACATTGTT	1575
OY	549	HisSerSerValAspArgTyrcAsnAlaIleSerAspLyvIleIleThrMetIle-----	566
Db	1576	GGCTTCTCACTTAG-----TTTCTTGATATTTTCCAGACCGGTGTTCCATGATATCATTA	1632
OY	567	---ProAlaIleLyvGIyAsnAsnLeuAspThrAsnSerLyvValIleGIuGIyProGIy	585

Db	1633	GCACCACTTATTAAGAAGAAAT-----TCCTAACCTTAGAGTGCTGAAAAGCAGCTCGA	1683
Oy	586	-----HisThrGlyGlyAsnLeuVal	592
Db	1684	TGCCGTAACTTGCTATTGTGAACATGATGGAAGTAGAAGCATGAGATGGTGGTTAATATATAC	1743
Oy	593	-----TyrLeuGlnSerGlnGlyArgLeuGluIleThrCysGluThrProAsnSerThr	610
Db	1744	GGAAGATTTTCTACAAA-----GAAATAAACAGTACTTGCTGGCTTAGAGAGATGTGAA	1794
Oy	611	GlnSerTyr-----PheIleArgLeu	617
Db	1795	ATGGATGGGCGATTTCGCCCTATTCGAGTTACCAACCTTATCTCTTTTAGTAAGTA	1854
Oy	618	ArgTyrAlaThrAsnGlyAlaGlyAsnThrLeuPro---AsnIleSerLeuThrIlePro	636
Db	1855	AGAAAGATGACTATGTGACCTTGACAACAATTACCAGAGATATATATGTGTGCTCTCCCT	1914
Oy	637	GlyVal	638
Db	1915	GATATT	1920
RESULT 11			
CNSO9YP2			
LOCUS			
DEFINITION	CNSO9YP2	2460 bp	mRNA linear HTC 06-FEB-2004
	Arabidopsis thaliana Full-length cDNA Complete sequence from clone		
	GSJLTPGH7ZA12 of Hormone Treated Callus of strain col-0 of		
	Arabidopsis thaliana (thale cress).		
ACCESSION	BX832541		
KEYWORDS	BX832541.1 GI:42455129		
SOURCE	HTC; GSJLT_cDNA.		
ORGANISM	Arabidopsis thaliana (thale cress)		
	Arabidopsis thaliana		
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;		
	Spermatophytes; Magnoliophyta; eudicotyledons; core eudicotyledons;		
	rosids; eurosoid II; Brassicales; Brassicaceae; Arabidopsie.		
REFERENCE	1 (bases 1 to 2460)		
AUTHORS	Castelli,V., Aury,J.M., Jallion,O., Wincker,P., Clepet,C.,		
	Menard,M., Cruaud,C., Quetier,F., Scarpelli,C., Schachter,V.,		
	Temple,G., Caboche,M., Weissenbach,J. and Salanoubat,M.		
	Whole Genome Sequence Comparisons and 'Full-length' cDNA Sequences:		
	A Combined Approach to Evaluate and Improve Arabidopsis Genome		
	Annotation		
JOURNAL	Unpublished		
REFERENCE	2 (bases 1 to 2460)		
AUTHORS	Genoscope.		
JOURNAL	Direct Submission		
	Submitted (18-NOV-2003) Genoscope - Centre National de Sequencage :		
	BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr		
	- Web : www.genoscope.cns.fr)		
	The sequences are based on single pass reads.		
	Life Technologies (a division of Invitrogen) members carried out		
	full-length libraries construction ; Temple G.		
	Genosope members carried out sequencing and annotation : Castelli		
	V., Aury J.M., Jallion O., Wincker P., Menard M., Cruaud C.,		
	Schachter V., Weissenbach J., Salanoubat M.		
	URGV INRA : Clepet C., Caboche M.		
	Annotation is based on the June 2003 version of the Arabidopsis		
	genome released by MIPS (Munich Information center for Protein		
	Sequences). 5 prime and 3 prime are assembled with Phrap.		
	http://www.genoscope.cns.fr/externe/sequences/Banque_projet_EF/Full		
	length		
	http://www.genoscope.cns.fr/cgi-bin/ggb?source=Arabidopsis.		
FEATURES			
Source	Location/Qualifiers		
	1..2460		
	/organism="Arabidopsis thaliana"		
	/mol_type="mRNA"		
	/db_xref="taxon:3702"		
	/clone="GSJLTPGH7ZA12"		
	/tissue_type="Hormone Treated Callus"		
	/ecotype="Col-0"		
	/plasmid="pCMVSPORT_6"		

gene 1..2460
 /gene="Ab5g20250"

ORIGIN

Alignment Scores:

Pred. No.:	0.397	Length:	2460
Score:	115.00	Matches:	170
Percent Similarity:	31.7%	Conservative:	109
Best Local Similarity:	19.3%	Mismatches:	269
Query Match:	3.0%	Indels:	332
DB:	4	Gaps:	49

US-10-783-417-2 (1-735) x CNS091P2 (1-2460)

```

QY 57 ThrGlnTYGlyAspAsnPhenGlu-----ThrPheAlaSerAlaAspThrIle 72
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 116 ACGAGGTACAGAGAAAACGAGAGAGAGAGATATGACATTAACCGCG----- 166
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 73 AlaAlaValSerAlaGlyThrIleVal-----SerGlyThrLeuAlaGlyIleGly 90
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 167 GTTCGATATTCGACGCGGAATCTAATCATCAAAAACCGGACGATCTAACCGGTGACCA 226
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 91 Gly-----LeuThrSerIleSer-----GlyProIle-----IleGly 103
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 227 GATTAATGTCATACGACGTCACGATCGGAGAGCTGACCGGTAGAAAGAGACTTTCGTCGA 286
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 104 AlaIle-----IleIleSerPheGlyThrLeuIleThrVal 115
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 287 GCTGATATTTAACAAGAGAGAGTAACACATCTGATCCGATCGCTTCGCAATTC 346
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 116 PheTyrProAlaGlyGlnIleAspIleThrValTTPThrGlnPheIleIleSerGlyIle 135
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 347 CCGTTATGTCCTGTTCCGTTTAAAG--CTCGGTGATGCGCCAGAGATGGAGAA 403
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 136 IlePheValAspThrPro-----LeuThrGluSer----- 145
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 404 ATGGGCGCAGATATTCCTTACAGACGCACTTATTTGTCGAGACCAAGATGGGTTC 463
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 146 -----IleIleGlnLeuIle 150
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 464 CACCTTGAGCTGACGAGCTAACCGCGTCGAGTGAACCAAGATTACACCGTTTC 523
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 151 LeuGlnThrLeuGlnGly--PheArgGlnIleLeuGln----- 162
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 524 TTGCCGTTATTCGAGAGATCTTCCGTTGTCGTCACAGAAAGTTAACGATGAGATT 583
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 163 -----SerTyrAsnThrAlaLeuAspAspTyrArgIleLeuIle 175
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 584 GAGCTCTGTTGGAGAGTGTGACGTCGACACTAAACGTCGCTTACTCACTCTCTG 643
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 176 ArgLeuGlnAlaProGlyLeuProPheSerAlaLeuGlnGlnAlaAlaLeuThrLeu 195
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 644 TATATTCACGCC--GGTACAGATCCGTTTCAGACATACGAGAGCTTATTCGACACTT 700
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 196 LysIleArgPheGlnAsnValHisAsnAspPheIleArgIleIleProGly----- 212
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 701 AAGTTCGATCTGAATAGTTCCTCGTCAACGTCACGAAAGAGACTTCACAGATCGTTGAC 760
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 213 ---PheGlnLeuGlnThrTyrIleThrLeuLeuProIleTyrAlaGlnAlaAlaAsn 231
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 761 TACTTCGATGAGTGTGACCTTGAGACGCGTTTATCAAGAGTGACTCAGGAAGCGCTGAA 820
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 232 PheHisLeuAsnLeuGlnGlnGlyAlaGlu-----LeuAlaAsp 245
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 821 GGTGATCTTAACTCTCGCCCGCGGTGTAGCCCGCAAGTTGTTATCATACACAC 880
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 246 GluTyrAsnAlaAspIleHisProSerGlnIleGlu----- 257
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 881 GGTGGCAATCAGTTGAGAGATGCGACGCTGAGGCGGAGATGAGAAAGAGTCA 940
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 258 -----Pro 258
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 941 CCGATTTTTCGACTGACGGGATCAAGAGAAAGAAAGTTTAAAGAGAGATGATCA 1000
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

```

```

QY 259 AsnAlaGlyThrSerAspAspTyrTyrIleLeuLeuGlu----- 272
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1001 AACGTTGGGATTAAG--AACATGTCAAGATGTCTAAAGAGAAAACCGTTGAAATAT 1057
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 272 ----- 272
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1058 GTTACGATGTCACGCGATACCGGTTATTTGGGCGGAGTTAGACCCGGGAGAAATAC 1117
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 273 ---AsnIleProIleTyrSerAsnTyrCysAlaAsn-----ThrTyr 285
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1118 GATCTCTGTATGAAATATCTTATATATGTCAAAAGGTGTGTGAGAAATGATCCGAGTGG 1177
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 286 ArgThrGlyLeuIleAsnLeuArgAsp-----GluProAsnMetLysTyrSer 301
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1178 AAGACTGATGATTAATACCGTCGACAGGTTGGGTTGTGTTAGCCCAAGAAAGTATATAG 1237
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 302 IlePheAsnAspTyrArgArgTyrMetThrIleThrValLeuAspThrIleSerGlnPhe 321
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1238 TTTTACAATGAGCTTCATATGTTACTTGTGCTGATGCTGGCGTGGACGCTGTGAAA----- 1291
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 322 SerLeuTyrAspIleIleAsnArgTyrArgAspSerIleGlyIleIleGluValIleGlyIle 341
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1292 -----GTGATGTCAAGTGTATTTGAGAGACTTGGGTGT-----GTTTACGCGGT 1339
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 342 LysAsnGlnLeuThrArgGlnIleTyr-----ThrThrGlnIleAsnPhe 356
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1340 CCGGTTGAGCTGACCTCGTCAGATTTTCATCAAGCTCTTGACTCTGTTGCTAAAGAACTTC 1399
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 357 AspArgLeuProGlnIleAsnArgValAlaProAsn-----LeuAlaThrMetGluTyr 373
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1400 -----CCGATTAATGCGTCGATGATTTGATGAGCAATGATGAT 1432
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 374 Asn-----LeuThrArgAlaSerPheLysLeu 382
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1433 AATACGATGCTCTTATTCGCTCGAAGCAAGAGCTGTGATTAAGACATCAATGAT 1489
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 383 PheSerPheLeuGlnGlnPheIlePheTyrThrGlnAsnThrAsnPheGlyAsnArgLeu 402
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1490 -----TTCTATCA----- 1498
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 403 ValGlyIleSerAsnArgAspAlaProThrTyrSerAsnThrIleThrGluThrLeuTyr 422
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1499 -----CGGATCCGGTGTCTCATACCATCAATACATGATCCCTGCTTAT 1543
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 423 -----GlyGlu-----ArgThrGlnSerProThrThrIleArg 435
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1544 AACAGTGTCTTGGAGAGATTATNGACAGCTGACCTGAGATATGTTCCATTCGTCGAC 1603
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 436 PropheGlnSerTyrIleValSerIleValThrAspArgGlnSerProProValSerPro 455
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1604 CCTGTGAGAGATATCATACGCTCTCTCTAGGCAATCATGCGGGCT----- 1651
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 456 IleGlnProHisPheIleIleAsnGlnIleGluLeuTyrLeuAsnGlySerSerAsnAsn 475
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1652 -----CTATGTCTAGGATTTCTCTGGA--- 1675
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 476 ThrLeuIleTyrSerAlaGlyIleSerLeuSerAsnTyrGlnAsnThrThrPheGln 495
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1676 ---AAGACACAC-----TTTGAG 1690
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 496 PheProArgLys----- 499
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1691 CTTCTAAGAAAGCTGTATGCTGTATGATGATGATTTCTTGCTGCACTCCCTGTAGA 1750
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 500 -----LysAspCysAsnLeuValIleAspPro-----GlyCysSerPro----- 512
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1751 CCAACTCCCGCATTTG---TTGTTCCGTGATCTTGCCCGCTGATGAGTGTGAGCTTGTAAAG 1807
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 513 ---AsnPheAsnAsnTyrSerHisIleLeuSerHisPheSerLeuPheThrTyrSer 530
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1808 ATATGAAACATGAAACAGTAC----- 1828
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

```


Db 570 GACATTGAGATCGCCACCTACAGAGCTGCGAAGGCGAGAGCGAGATTCTCTCG 629
Qy 361 LysleuPheSerPheleu-----GluGlnPheIlePheIleThrGlnAnthr 396
Db 630 CCTCTGCCACCTTTTCTTCTCCGAACTTGAAGAACTACTAGATCCTCTCTCT 683
Qy 397 AenPheGlyAenArgleuValGlyIleSerAenArgAspAlaPro 411
Db 684 -----CTGGTTGACACCACTCANAAAGACATCTCT 713

RESULT 13
AG378595/c 1123 bp DNA linear GSS 21-DEC-2004
LOCUS Mus musculus molossinus DNA, clone:MSMg01-187C04.TU, genomic survey
DEFINITION sequence.
ACCESSION AG378595
VERSION AG378595.1 GI:47989800
KEYWORDS GSS.
SOURCE Mus musculus molossinus (Japanese wild mouse)
ORGANISM Mus musculus molossinus
Bukeryoca; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1
Abe,K., Noguchi,H., Tagawa,K., Yuzurika,M., Toyoda,A., Kojima,T.,
Ezawa,K., Satou,N., Hattori,M., Sakaki,Y., Moriwaki,K. and
Shiohama,T.
TITLE Contribution of Asian mouse subspecies Mus musculus molossinus to
genomic constitution of strain C57BL/6J, as defined by BAC-end
sequence-SNP analysis
JOURNAL Genome Res. 14 (12), 2439-2447 (2004)
PUBMED 15574823
AUTHORS Hattori,M., Toyoda,A., Noguchi,H., Kojima,T. and Sakaki,Y.
TITLE Direct Substasion
JOURNAL Submitted (17-NOV-2003) Masahira Hattori, The Institute of Physical
and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
1-7-22 Suenhiro-chou,Tsuri-mi-ku, Yokohama, Kanagawa, 230-0045, Japan
(E-mail:hattori@gsc.riken.jp, URL:http://hgp.gsc.riken.go.jp/,
Tel:81-45-503-9111, Fax:81-45-503-9170)
COMMENT Clones are derived from the mouse BAC library MSMg01. For BAC
library availability, please contact Kunya Abe (abe@rtc.riken.jp).
Tsukuba Institute, Bio Resource Center,
The Institute of Physical and Chemical Research (RIKEN) 3-1-1
Koydai, Tsukuba, 305-0074 Japan
phone: 81-298-36-9189, fax: 81-298-36-9199
e-mail: abe@rtc.riken.jp
PRIMERS
Sequencing : TU
LIBRARY
Vector : PBAC3.6
R.Site 1 : EcoRI
R.Site 2 : EcoRI.

FEATURES
source location/Qualifiers
1..1123
/organism="Mus musculus molossinus"
/mol_type="genomic DNA"
/sub_species="molossinus"
/db_xref="taxon:57486"
/clone="MSMg01-187C04.TU"
/sex="male"
/tissue_type="mixture of kidney and spleen"
/clone_lib="MSMg01 Mouse Male BAC library"

ORIGIN
Alignment Scores:
Pred. No.: 0.218 Length: 1123
Score: 112.50 Matches: 97
Percent Similarity: 38.04 Conservative: 53
Best Local Similarity: 24.64 Mismatches: 147
Query Match: 2.94 Indels: 99
Db: 10 Gaps: 19

US-10-783-417-2 (1-735) x AG378595 (1-1123)
Qy 352 ThrGluIleAenPheAspArgleuProGlnleuArgValGlnProAsnleuAlaThrMet 371
Db 1081 ACCCCCAATAACACACCTTTCTTCCCAAACTACTCTCTTGTATATATATACACCTTC 1022
Qy 372 GlnTYAenleuThrArgAlaSerPheIleuPheSerPheleuGlnIlePheIlePhe 391
Db 1021 TCTTATATCTTACACCACT-----CTCTCTCTTCAACAACCAAAATTAACAACC 971
Qy 392 Tyr-----ThrGluAenThrAenPheGlyAenArgleuValGlyIle 405
Db 970 TATTTTCATTAATACACACCAATCCGAAACACCACTACCTCATTTAATATATATCTTC 911
Qy 406 SerAenArgAspAlaProThrTYrSerAenThrIleThrGlnleuTYrGlyIleArg 425
Db 910 CAACAT-----ACCTACACCACTACTCTACATCAACATCAACACACAACTT 863
Qy 426 ThrGlySerProThrThryleThrIleArgProPheGluSerTYrIleValSerIleVal 445
Db 862 CCTATTTCAATTAACAAACCTAGAAACAATATATCTCCCTCTTATCAATATCTTC 803
Qy 446 ThrAspArgGlnSerProProValSerProIleGlnProIlePheIleleuGlnIle 465
Db 802 -----TCCCAACCAATCAATCAACACACACCCCTTTTTCATCTCAATATATC 755
Qy 466 GlnleuTYrleuAsnGlySerSerAenAenThrleuTYrSerAlaGlyIleSerleu 485
Db 754 TCATTAATACAAATATCCCTCTT-----TTAAATATCCCTCTTATATCTTAT 704
Qy 486 SerAenTYrGlnAenThrThryPheGlnPheProArgIleValAspCysAenleuVal 505
Db 703 CTCATATCTCTTCCATATCATCTTACAA----- 674
Qy 506 IleAspProGlyCysSerProAenPheAenAenTYrSerHis-IleleuSerHisPhe 525
Db 673 CTCATATCAACCTCTCTCCCT-----TCCACATCACTACTTATATCTTAT 629
Qy 525 ThrleuThrTYrSerTYrValIleGlyleuGlnleuGlnIle----- 539
Db 628 TTTTCT-----TTTTGT 575
Qy 540 -leuAspThrGlyValleuGlyTYrPheHisSerSerValAspArgTYrAsnAlaIle 559
Db 574 CATATCCCAACCAAAACCTTTCTTGTCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 515
Qy 555 rAspIleIleIleThrMetIleProAlaIleuGlyAenAenleuAspThrAsnSerTY 579
Db 514 TCTTCTCACTCTCTCTCCACCCCTCCCTTAAAAAAATCTT-----ACATATCTCT 461
Qy 579 eValIleGlnGlyProGlyIleThrGlyIleValleuValTYrleuGlnSerGlnIle 599
Db 460 TATTAATTAATTTCTT-----TTCTTCACTCTG----- 433
Qy 599 GlnleuIleThrCysGluThrPro-----AsnSerThrglnSerTYrPheIle 615
Db 432 -----CCCAACCCCTCTCTCTATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 391
Qy 615 eArgleuArgTYrAlaThrAsnGlyAlaGlyAenThrleu-----ProAsnIle 632
Db 390 -----ACCACTAAATAAATCTCTTCTTCTCTCTCTCTCTCTCTCTCTCTCTCT 347
Qy 632 eTleuThrIleProGlyValIleGlyIleProProGlnArgleuAenAenThrPheSer 652
Db 346 CTCTTCACTCTCAATTAATCTCTTACCTCCCTCTCTCTCTCTCTCTCTCTCTCTCT 287
Qy 652 Ty-----ThryAenTYrAenleuGlnTYrGlyAspPheGlyTYrPheGlnP 668
Db 286 CTTTTAAACATATCTCTCTCAATATCAATTAATATATTTTCTTCTCTCTCTCTCT 227
Qy 668 hePro-----SerThryValThryleuProleuAenArgAenIle----- 680

Db	226	ACCCCTTTTGGGAAAAAACAACCTCTCCCTTCACTA-AACTACTTCTTT	168
Oy	681	-----ProPheilePheAsnArgAlaAspValSerAsnSerileu	695
Db	167	AATAAAACCGCCACCTTTTGTG-----AAATAATCCTTCTCC	123
Oy	695	leileAspyleleleluPheilePheileTrSerSerMet	708
Db	122	GCCCTGCGCCCTTTTATTTTACACCCCAATTACTCTCTCC	82
RESULT 14			
CL974457		2841 bp	DNA
LOCUS			linear
DEFINITION			GSS 21-SEP-2004
OsJFCC025896			Oryza sativa Expressed Library Oryza sativa (indica
CL974457			cultivar-group) genomic, genomic library sequence.
CL974457.1			GI:52403441
KEYWORDS			GSS.
ORGANISM			Oryza sativa (indica cultivar-group)
REFERENCE			Oryza sativa (indica cultivar-group)
AUTHORS			Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
			Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
			Ehhartridae; Oryzaceae; Oryza.
TITLE			1 (bases 1 to 2841)
JOURNAL			Ma, L., Wang, J., Chen, C., Liu, X., Su, N., Li, L., Wang, X., Cao, M.,
COMMENT			Jiao, Y., Sun, N., Zhang, X., Bao, J., Sun, D., Zhao, H., Yuan, L.,
			Wong, G. K. S., Deng, X. W. and Wang, J.
			An analysis of transcriptional regulation of the rice genome and
			its comparison to Arabidopsis
			Unpublished (2004)
			Contact: Chen Chen
			Department of Bioinformatic
			Beijing Institute of Genomics
			Chinese Academy of Sciences, Beijing 101300, China
			Tel: 86-10-80481559
			Fax: 86-10-80488676
			Email: chenchen@genomics.org.cn
			Rice genomic sequence.
			Class: exon-trapped.
FEATURES			Location/Qualifiers
source			1..2841
			/organism="Oryza sativa (indica cultivar-group)"
			/mol_type="genomic DNA"
			/db_xref="taxon:39946"
			/clone_lib="Oryza sativa Expressed Library"
			/note="Oryza sativa exon trapped genomic sequences "
ORIGIN			
Alignment Scores:			
Pred. No.:	0.982	Length:	2841
Score:	112.50	Matches:	175
Percent Similarity:	33.1%	Conservative:	111
Best Local Similarity:	20.2%	Mismatches:	313
Query Match:	2.9%	Indels:	267
DB:	10	Gaps:	47
US-10-783-417-2 (1-735) x CL974457 (1-2841)			
Oy	24	ArghAsnSerAsnAspSerArgTyProTyThrAsnAsnProAsnGln-----Proleu	41
Db	499	CGCAATGAAGCGTTGCAAAAGAAATTTGCACACATTAACCAATVAGGCTAAATCCCTG	558
Oy	42	GlnAsnThrAsnTyLysGluTrpLeuAsnMetCysGlnGlnYAsnThrGlnTyrclyAsp	61
Db	559	GACTTACC-----ATGTTTGAAATATATATTAAGCAACGACGT	597
Oy	62	AsnPheGluTrpPheAlaSerAlaAspThrIleAlaValaSerAlaGlyThrIleVal	81
Db	598	AATTAGAGCACTAGACGTGATTAAGACCAATTCAGCAAGTC-----GATGGGTCAATT	651
Oy	82	SerGlyThrLeuLeuAlaGlyTlGlyGlyLeuThrSerIleSerGlyProIleGlyTle	101
Db	652	GCC-----AGAAATGTGAGAGCAATVAGCATTAAGACGCCCACTGGTCTTC	699

OY	102	11e-----g1yallellellelserp	heglythrleuilethrvalphetrProala	119
		:::		
Db	700	TTGCTATCTGGAAATGACCTTGAGTAATGACAGACATCG	-----	738
OY	120	g1ygluglnabpbylsthValtrpThrc1nphel	lelyswetg1ygluilepheValasp	139
		:::		
Db	739	-----AGCAATAGTGAATATTGAACT	-----	774
		:::		
OY	140	ThrProleuthrnguseSerlelelyeglInleulyseuglnthrleuglnlyphearg3ln	159	
		:::		
Db	775	AATTCAAGTGGTCCATCCTTGAAA-----CTAACTTACAAGAAATATGCCACCATATTGAGA	831	
OY	160	11leleuglnserlyr-----	Asnthrala-----	167
		:::		
Db	832	TTAATCTTTGGCTTACCTGCGCCATCTTCCCTTAAGAGTCAACAATATATAGCTTAAAGCAAGCTCG	891	
OY	168	LeuaspasbpTTParglylyleulyeArgheuglnlaProglyleuProPserSerla	187	
		:::		
Db	892	GTTCAACCAATGATGCTCTTGTTTATTATGAG-----	CCATCTTAAACA	936
OY	188	LeuGlnglnla1a1aleuthrleulyel1earPhegluAenValh1asnaAspheile	207	
		:::		
Db	937	TTTTATCCCGTACGACCTTGGGAGAAAG-----	TATATC	969
OY	208	ArggluileProglylypheglInleuglnuthrlyrlysthThrleuleuProileTyra	227	
		:::		
Db	970	AGACAACATCTTTGGG-----ATgncTTCCTTCAGCGTTCATATTCACATGACATGACAA	1022	
OY	228	Glnla1a1aasnpheh1seuasnleuenglnnglnlylaagluAul1aasp1uTrp	247	
		:::		
Db	1024	GAACTTTTTCACACATGACGATATGATGTTCTATGATGTTGGAGATCGCTATGAGTAAGAAAG	1083	
OY	248	Asnla1aaspl1e1sPserGln11egluProAsnla1ag1ythrSerAspaspTyrr	267	
		:::		
Db	1084	TTGGTTTTCTTTAATATGACACAAAATTTATGACGTACACACAGAACAAAATTTTGCACATAT	1143	
OY	268	LyseuleulysgluAenlleProlyseTySerAsnTyrr-----	280	
		:::		
Db	1144	GCTTTGGCTT-----GAAATATATAGCAAGTCATCAATCATATCACTATATTGCCGCCACG	1200	
OY	281	-----	Cye1a1aasnthrTyraArgthrglyr-----	288
		:::		
Db	1201	TTAAGGCGAGTACATACTTCGAAATTCGACGAAATTCGGTGCCTCAAGGATGAATATTTCA	1260	
OY	288	-----	-----	288
Db	1261	TTTCAAAAGTTCTTGCGTGTCTTAGAATTTAACTGACTGCTCCATAGAGATCTCCCAAGC	1320	
OY	289	-----LeulyAsenleuArg-----	Asp1uProAsnMetyleTrrSerile	302
		:::		
Db	1321	TCTATATGGCAATTTGAAAGCACTGAAGCTTTCTTATCGCTCAAAATCATGAGATATATATG	1380	
OY	303	PheasnaspTyrraArgaArgyrmethThr1lethrValleuAspThr1leSerGlnPheSer	322	
		:::		
Db	1381	TTCCCA-----	AAAGATATTACTGTGCTCCCAAAACTA-----	1413
OY	323	LeutyrrAsp1leulyAsArgTyrrArgasbpserlleelyglylegluVallysglylelys	342	
		:::		
Db	1414	-----AAATACCTTGACCTCCATGAGATCTTTTGCAAATTTTCAGATTTGCA	1458	
OY	343	AsngluethrThraArggluileTyrrThrThrc1u1leasnPheasbpArgleuProglIneu	362	
		:::		
Db	1459	GGATCAATATGCCAAACATGCTTGCTGTGATACATCTTGACTTATCGTTCCTTAAATATC	1518	
OY	363	ArgVal---GlnProAsnleuAlaThrMetGluTyrrAsnleuthrThraArgalaserPhelys	381	
		:::		
Db	1519	AGAAGTTATACAGCCCAAGACTCTAATGT-----GGCCTCACCAAACTGCAGATTTTG	1566	
OY	382	LeuPheSerPheleuglnInPhe1lePheTyrrThrc1u1asnthrAsnphelg1yAsnArg	401	
		:::		
Db	1570	AACTTATCTGTGCTCCATTTCTTCAATAATCTGCAGAAAATATTTGCAAGCTTACTGAA	1629	

OY	402	Leu-----ValGlyIleSerAsn-----ArgAspAlaProThrTySerAsn	415
Db	1630	CTGCACAAATTAAACTTATACAAACCTGCTCTCTCAGCAGCTTACCACTTCATATTTGGG	1689
OY	416	ThrIleThrgIuThrLeuTyrgIyGluArgThrgIy---SerProThrThrLysThrIle	434
Db	1690	AGCCTCACTGAACCTACATATTTGAACTTATACAGGTGTCAAGGGCTTGTAAATTCGCA	1749
OY	435	ArgProPheGluSerTyrlLysValSerIle-----	444
Db	1750	ATGTGCTTTCGGAATCTGAAGAAATTGGTACATCTTGACCTGTGAGGCTGTCCGGGTA	1809
OY	445	-----ValThrAspArgGlnSerProProValSerPro	455
Db	1810	CAAGATTTCAAACAAGCTTTGTTGGTGGCTACCAAACTGCAGTACTTAATCTTAACAAAG	1869
OY	456	Ile-----GlnProHisPheIle-----	461
Db	1870	ATATTGGTAGAACACAGATGGGTGTATATTGGGATGGTATTCTCGAAACCTATCAGTACT	1929
OY	462	IleAsnGlnIleGlnLeuTyrlLeuAsnGlySerSerAsnAsnThrLeuLysTyrl-----	479
Db	1930	CTTAATGATCTCTTGAG---TATCTAAACCTTCTCGGAATACAGAGATTGATTAATCTACA	1986
OY	480	SerIAsnGlyIySerLeuSerAsnTyrgIeAsnThrThrPhePheGlnPheProArgLys	499
Db	1987	AGAACCTTTGGAAACCTCAAAAAGCTTCAA---ACATTGAACTTTCAATACGTGTGGCT	2043
OY	500	-----LysAspCysAsnLeuValIleAspPro	508
Db	2044	CTGAGATCCCTCCACACAGCATAGAAATTAATGACATCTCGAGTTCTATATGTTGTG	2103
OY	509	GlyCysSerProAsnPheAsnAsnTyrl-----Ser	518
Db	2104	GGGTGCTCAGATCACTTAAGGAGTATTGGAAATCTCATTCAGAATATTCAGTCT	2163
OY	519	HisIleLeuSerHisPheSerLeuPheThrTyrlSerTyrlValIle-----Gly	534
Db	2164	CATTAC---CACACTTCACTTGTCACTCATATTTCTGCCATAGAAAGAACAGAGGA	2217
OY	535	LeuGlnLeu-----GlnIleLeuAspThrgIyValIleGlyTyrlThrHisSerSer	551
Db	2218	ATTGAGTTGTCAAGAAACAAAATCTTTCACATTTGACATTTCAATGAGACT---TCAAG	2274
OY	552	ValAspArgTyrlAsnAlaIleSerAspLysIleIleThrMetIlePro-----	567
Db	2275	GCTGACCGACTTTTGAGAGACACAGATGTTTGGAGAACTAATGCCACCAAGAGTCTG	2334
OY	568	-----AlaIleLysGlyAsnAsnLeuAspThrAsnSerLysValIleGlnIyPro	584
Db	2335	TGGAATTTAAGCATACAAAGGATATATGACACAACTTCCCAACTGTGTTGTAGGATTC	2394
OY	585	GlyHisThrThrgIyGlyAsnLeuValTyrlLeuGlnSerGlnIyArgLeuGlnIleThrCys	604
Db	2395	TCACATCATCTTCCAAAGCTTGTTAAGATGAGTT- GAAGATTTGCCACAATG-----	2447
OY	605	GluThrProAsnSerThr-----	610
Db	2448	CAGACACCTGCACACCATCGGACAGTTCGCAAACTGATGATGATGATCTCCAAAGAT	2507
OY	611	---GlnSerTyrl-----PheIleArgLeuArgTyrlAlaThrAsnGlyAla	624
Db	2508	GGACAGCTCTCAAAATTAATGATGGCGGTTTCTGCGGTGGCCACA-----	2552
OY	625	GlyAsnThrLeuProAsnIle---SerLeuThrIleProGlyValIleGlyIleProPro	643
Db	2553	---AGCGCATTCGCCAAACTGAAGATCTTCACTTGAGGGATAT-----	2594
OY	644	GlnArgLeuAsnAsnThrPheSerGlyThrAsnTyrlAsnAsnLeuGln-TyrgIy-----	661
Db	2595	-----GAAACAGTTGGAAGATGGAGACACCAACACACATTAAGATGTGTA	2642
OY	662	AspPheGlyTyrlPheGlnPheProSerThrValThrLeuProLeuAsn-----	677

Df		2643	CGATAGCATAGATTCACTTGTTCCTATGCTACGTACTGCATACAGCTTTTCCAA	2702
Oy	678	-----ArganilleProphellpeAsnaraAlaspValseranserrillele	694	
Df	2703	GCTTAGACTGAACAACGATGCCACCGATGTTTCACAGCTGGCTCATATCAGAGACGACAA	2762	
Oy	694	uuilaleasp-----lytle-----GiuphelllePr	703	
Df	2763	GAGTGTCACACGGGCTTTAACGAGAGCTTGATGATAATTTTCCGTGAAGATTTTCATGCC	2822	
Oy	703	oileThrseser 707		
Df	2823	TACCGACTTTCA 2835		
RESULT 15				
CNSOALJUE				
LOCUS				
DEFINITION				
ACCESSION				
VERSION				
KEYWORDS				
SOURCE				
ORGANISM				
REFERENCE				
AUTHORS				
TITLE				
JOURNAL				
REFERENCE				
AUTHORS				
TITLE				
JOURNAL				
COMMENT				
FEATURES				
source				
ORIGIN				
Alignment Scores:				
Pred. No.:				
Score:				
Percent Similarity:				

THIS PAGE BLANK (USPTO)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: January 20, 2006, 14:10:00 ; Search time 1091 Seconds
(without alignments)
4489.962 Million cell updates/sec

Title: US-10-783-417-2
Perfect score: 3869
Sequence: 1 MNQNNNDNNYEIIIDSHSPY.....KLETIQTKNPFTNHTKTL 735

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 4996997 seqs, 3332346308 residues
Total number of hits satisfying chosen parameters: 9993994

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame+gln.model -DEV=Xlh
-O=/cgn2.1/USPTO_epool/US10783417/runat_20012006_095009_23859/app_query.fasta_1.903
-DB=N Geneseg -OFMT=fastcap -SUFFIX=p2n.rng -MINMATCH=0.1 -LOOPEXT=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blomsum62 -TRANS=human40.cdi -LIST=45
-DOALIGN=200 -THR_SCORE=pcr -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pcr -NORM=ext -HEAPSIZE=500 -MILEN=0 -MAXLEN=2000000000
-USER=US10783417@cgn2.1 752 @runat_20012006_095009_23859 -NCPU=6 -ICPU=3
-NO MMAP -NEG SCORES=0 -WAIT -DSPBLOCK=100 -DEV TIMEOUT=120 -WARN TIMEOUT=30
-THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -Fgapop=6 -Fgapext=7 -Ygapop=10
-YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

N_Geneseg_21.*
1: genesegn1980s:*
2: genesegn1990s:*
3: genesegn2000s:*
4: genesegn2001as:*
5: genesegn2001bs:*
6: genesegn2002as:*
7: genesegn2002bs:*
8: genesegn2003as:*
9: genesegn2003bs:*
10: genesegn2003cs:*
11: genesegn2003ds:*
12: genesegn2004as:*
13: genesegn2004bs:*
14: genesegn2005s:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3869	100.0	2208	13	ADR89394 AXMI-006
2	3314	85.7	2235	13	ADR89396 AXMI-007
3	3104.5	80.2	2085	13	ADR89398
4	1061.5	27.4	4571	1	AAN93059 Delta-end

5	1060.5	27.4	3543	2	AAQ14669	AAQ14669 Dipteran
6	1060.5	27.4	3543	2	AAQ81178	AAQ81178 B.c. coxi
7	1059.5	27.4	3940	2	AAN93054	AAN93054 Delta-end
8	1059.5	27.4	4934	1	AAN81490	AAN81490 Insectici
9	1032.5	26.7	2130	12	ADP71294	ADP71294 Codon-mod
10	1032.5	26.7	2246	12	ADP71295	ADP71295 Codon-mod
11	899.5	23.2	2061	2	AAQ14670	AAQ14670 Dipteran
12	899.5	23.2	2061	2	AAQ81180	AAQ81180 B.c. coxi
13	881	22.8	3756	1	AAN50525	AAN50525 Bacillus
14	849	21.9	4896	6	AAQ43974	AAQ43974 Bacillus
15	849	21.9	4896	10	ADP31301	ADP31301 Bacillus
16	849	21.9	4896	10	ADP31306	ADP31306 Bacillus
17	775.5	20.0	3507	2	AAQ43221	AAQ43221 Anticarsa
18	759	19.6	3684	6	ABK14949	ABK14949 Bacillus
19	759	19.6	3684	14	AEA61392	AEA61392 Bacillus
20	758	19.6	3684	1	AAN82106	AAN82106 Sequence
21	757.5	19.6	3535	1	AAN91003	AAN91003 Sequence
22	747.5	19.3	4184	1	AAN96139	AAN96139 Bacillus
23	711.5	18.4	4065	1	AAN92515	AAN92515 Sequence
24	705	18.2	4391	14	AEA61394	AEA61394 Bacillus
25	703	18.2	3633	6	ABK87235	ABK87235 Bacillus
26	703	18.2	3633	12	ADL15306	ADL15306 B. thuring
27	703	18.2	3633	14	ABK45608	ABK45608 B. thurin
28	703	18.2	6613	6	ABK87248	ABK87248 Bacillus
29	703	18.2	6613	12	ADL15321	ADL15321 B. thuring
30	703	18.2	6613	14	ABK45623	ABK45623 B. thuring
31	696.5	18.0	10974	1	AAN60055	AAN60055 VB131 c10
32	689.5	17.8	4188	8	ABK13487	ABK13487 B. popill
33	689.5	17.8	4188	14	ABK90816	ABK90816 Paenibaci
34	678.5	17.5	2003	6	ABK87237	ABK87237 Bacillus
35	678.5	17.5	2022	6	ABK87242	ABK87242 Bacillus
36	678.5	17.5	2022	12	ADL15316	ADL15316 B. thuring
37	678.5	17.5	2022	14	ABK45618	ABK45618 B. thurin
38	677	17.5	2478	12	ADP43079	ADP43079 Bacillus
39	676	17.5	1953	2	AAQ43223	AAQ43223 Anticarsa
40	676	17.5	1953	3	AAQ09499	AAQ09499 Plant-cop
41	676	17.5	2425	2	AAQ05679	AAQ05679 bPGS1208
42	675	17.4	3621	6	ABK87234	ABK87234 Bacillus
43	675	17.4	3621	12	ADL15304	ADL15304 B. thuring
44	675	17.4	3621	14	ABK45606	ABK45606 B. thurin
45	675	17.4	3621	14	ABK34684	ABK34684 Bacillus

ALIGNMENTS

RESULT 1
ADR89394
ID ADR89394 strand; cDNA; 2208 BP.
XX
AC ADR89394;
XX
DT 18-NOV-2004 (first entry)
XX
DE AXMI-006 coding sequence.
XX
KW ss; gene; delta-endotoxin; delta-endotoxin associate polypeptide;
KW expression cassette; transfection; transgenic; plant; bacteria;
KW lepidoptera; coleoptera; pest; pesticide; resistance;
KW pesticial activity.
XX
OS Bacillus thuringiensis.
FH
FT Key Location/Qualifiers
FT CDS 1..2208
FT /tag= a
FT /product= "AXMI-006"
XX
PN MO2004074462-A2.
XX
PD 02-SEP-2004.
XX
PF 20-FEB-2004; 2004WO-US005829.
XX


```

Db 1381 ATAAATTAATCAAAATGGAATTATTAATTAATGCTCATCAACACACTCAAAATATCA 1440
QY ALaGIyGlySerLeuSerAsnTyrGlnAsnThrPhePheGlnPheProArgLys 500
Db 1441 GAGAGAGGCTCTTAATCAATATATCAAAACACAACTTTTTCATTTCTTGAAAANA 1500
QY AspCysAsnLeuValIleAspProGlyCysSerProAsnPheAsnTyrSerHisIle 520
Db 1501 GACTGCATCTAGTATTAATGATCCAGTTGTTCAACAACTTAAATACATATGATATTT 1560
QY LeuSerHisPheSerLeuPheThrTyrSerTyrValIleGlyLeuGlnLeuGlnIleLeu 540
Db 1561 TTATCCCATTTTTCATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1620
QY AspThrGlyValIleGlyTyrThrHisSerSerValAspArgTyrAsnAlaIleSerSer 560
Db 1621 GATACAGGTGTATGATGAGACACACAGTACTGTGTATGATATGATATGATATCAAT 1680
QY LysIleIleThrMetIleProAlaIleLysGlyAsnAsnLeuAspThrAsnSerLysVal 580
Db 1681 AAAATTAATTAATCAATATCCAGCATCAAGGTAACATCTTGATACAAACTCTAAGGA 1740
QY IleGlyGlyProGlyValIleThrGlyValAsnLeuValTyrLeuGlnSerGlnGlyArgLeu 600
Db 1741 ATTGAAGACCTGTCATACAGAGGAAACTGGTTTATTAACAAGTCAAGGGGTTTA 1800
QY GluIleThrCysGluThrProAsnSerThrGlnSerTyrPheIleArgLeuArgTyrAla 620
Db 1801 GAAATTAATCAATGTAAGAACTCTTAATTTCAACAATCTTAATTCATTAAGACTGATATGCT 1860
QY ThrAsnGlyValIleGlyAsnThrLeuProAsnIleSerLeuThrIleProGlyValIleGly 640
Db 1861 ACAAATGGTCTGGAATTAATCTTCTTAATATATCTTACAAATCCAGGATTAATGGA 1920
QY IleProGlnIleArgLeuAsnAsnAsnThrPheSerGlyThrAsnTyrAsnAsnLeuGlnTyr 660
Db 1921 ATACACCTCAACGACTCAACAACTTTTCTGATACAAATTAATTAATTAATTAATCAATAC 1980
QY GlyAspPheGlyTyrPheGlnPheProSerThrValIleThrLeuProLeuAsnArgAsnIle 680
Db 1981 GGAGATTTTGGGTATTTCCAAATTTCCAAATGACATTAATCTTAAATGAAACATA 2040
QY ProPheIlePheAsnArgAlaAspValSerAsnSerIleLeuIleIleAspValIleGlu 700
Db 2041 CCATTAAATTAATTAATCGTCAGATGTATCAAAATTAATTAATTAATTAATTAATTAATGA 2100
QY PheIleProIleThrSerSerMetHisGlnAsnArgGluLysGlnLysLeuGlnThrIle 720
Db 2101 TTTATACCAATTAATCTTCTCTATGACCAAAATAGGAAAAACAAAATTAAGAAACTATC 2160
QY GlnThrLysIleAsnThrPhePheThrAsnHisSerThrLysThrLeu 735
Db 2161 CAAACAAAATTAATTAATTAATTTTTCACAAATCATTAACAAAACACTT 2205

```

```

RESULT 2
ADDR89396
ID ADDR89396 standard; cDNA; 2235 BP.
AC ADDR89396;
XX 18-NOV-2004 (first entry)
DT 18-NOV-2004 (first entry)
DE AXMI-007 coding sequence.
XX ss: gene; delta-endotoxin; delta-endotoxin associate polypeptide;
XX expression cassette; transformation; transgenic; plant; bacteria;
KM lepidoptera; coleoptera; pest; pesticide; resistance;
XX pesticidal activity.
OS Bacillus thuringiensis.
FH Key location/Qualifiers

```

```

FT CDS 1..2235
FT /tag= a
FT /product= "AXMI-007"
FT /trans_except= pos:1..3, aa:Met
FN MO2004074462-A2.
PD 02-SEP-2004.
XX 20-FEB-2004; 2004MO-US005829.
XX 20-FEB-2003; 2003US-0448632P.
XX 20-FEB-2003; 2003US-0448633P.
XX 20-FEB-2003; 2003US-0448797P.
XX 20-FEB-2003; 2003US-0448806P.
XX 20-FEB-2003; 2003US-0448812P.
XX 20-FEB-2003; 2003US-0448812P.
XX 19-FEB-2004; 2004US-00781979.
XX 19-FEB-2004; 2004US-00782020.
XX 19-FEB-2004; 2004US-00782096.
XX 19-FEB-2004; 2004US-00782141.
XX 19-FEB-2004; 2004US-00782570.
XX 19-FEB-2004; 2004US-00783417.
XX (ATHE-) ATHENIX CORP.
PI Carozzi N, Hargis T, Koziel MG, Duck NB, Carr B;
XX WPI; 2004-635574/61.
DR P-PSDB; ADR89397.
XX
PT New isolated delta-endotoxin and delta-endotoxin-associated nucleic acids
PT and polypeptides; useful for killing lepidopteran or coleopteran pests or
PT for producing organisms with pesticide resistance.
XX
PS Claim 1; SEQ ID NO 8; 178bp; English.
XX
CC This sequence encodes an isolated delta-endotoxin. Some of the delta-
CC endotoxin coding sequences of the invention have alternative start
CC codons, producing more than one protein from a single open reading frame.
CC The nucleic acid sequences of the invention are useful in DNA constructs
CC or expression cassettes for transformation and expression in plants and
CC bacteria. The nucleic acids and corresponding polypeptides are useful for
CC killing lepidopteran or coleopteran pests. Compositions containing the
CC delta-endotoxins of the invention, and methods for their production, are
CC useful for the production of organisms with pesticide resistance.
CC specifically bacteria and plants. These organisms are useful for
CC generating altered or improved delta-endotoxin or delta-endotoxin-
CC associated proteins that have pesticidal activity, or for detecting the
CC presence of delta-endotoxin or delta-endotoxin-associated proteins or
CC nucleic acids in products or organisms.
SQ Sequence 2235 BP; 861 A; 364 C; 316 G; 694 T; 0 U; 0 Other;
XX
XX
Alignment Scores:
Pred. No.: 9,256-290 Length: 2235
Score: 3314.00 Matches: 638
Percent Similarity: 91.34 Conservative: 32
Best Local Similarity: 86.94 Mismatches: 60
Query Match: 85.74 Indels: 4
DB: 13 Gaps: 3
US-10-783-417-2 (1-735) x ADR89396 (1-2235)
QY 1 MetAsnGlnAsnAspAsnAsnGluTyrGluIleIleAspSerHisThrSerProTyr 20
Db 1 GTGATCAAAATTAAT---AATATGATATATGATATCGATTCAAAGAAATTATCTAT 57
QY 21 PheProAsnArgAsnSerAsnAspSerArgTyrProTyrThrAsnAsnProAsnGlnPro 40
Db 58 CCTTCTAAACAGAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATCA 117
QY 41 LeuGlnAsnThrAsnTyrLysGluThrPheLeuAsnMetCysGlnGlyAsnThrGlnTyrGly 60

```


Db 118 TTACAAAACAAAATACAAAAGAGGCTCAATATGTCAGAGGAAATACCAATATGCT 177
Qy AspaenPheGluThrPheAlaSerAlaAspThrIleAlaValaSerAlaGlyThrIle 80
Db 178 GATATTTTCGAGACATTTGCTAGTCGTACATATTCGTCAGTTAGTCGAGTACTATT 237
Qy 81 ValSerGlyThrLeuLeuAlaGlyIleGlyGlyLeuThrSerIleSerGlyProIleGly 100
Db 238 GTATCCGGTACTCTGTAGCCGGTATAGTGCGCTCACTTCTATATCCGAGCCGATAGGA 297
Qy 101 IleIleGlyAlaIleIleIleSerPheGlyThrIleIleThrValPheThrProAlaGly 120
Db 298 ATAAATAGGTCTATATATATATCTTTGGTACCTCAATACACGTCTTTGGCCCGGGA 357
Qy 121 GluGlnAspLeuThrValIleThrGlnPheIleIleMetGlyGlyIlePheValaAspThr 140
Db 358 GACACAGACAAAACAGTATGACACAAATTTATTTAAATGGAGAAATTTTGTGATAC 417
Qy 141 ProLeuThrGluSerIleGlyGlnLeuGlyLeuGlnThrLeuGlyPheArgGlnIle 160
Db 418 CCGTTAAACGAAAGCATAAACAGCTAAAGTTACAACTTTAGAAAGATTAGACAAATA 477
Qy 161 LeuGlnSerTyrrAsnThrAlaLeuAspAspTyrParglySLeuLyArgLeuGlnAlaPro 180
Db 478 TTACAAAGCTATAATACAGCATTTAGATGATGAGAAATTTAAAGACTACCAAGCTCCT 537
Qy 181 GlyLeuProProSerSerAlaLeuGlnGlnAlaAlaLeuThrLeuIleIleArgPheGly 200
Db 538 GATTAACACACATCATCGCATTTACACAGCTGCTTGACTTTAAATACATGATTTGAG 597
Qy 201 AsnValIleAsnAspPheIleArgGlnIleProGlyPheGlnLeuGlnThrTyrrAspThr 220
Db 598 AATGTTCAACAATGATTTTATTCGAGAAATACCTGGTTCCAACTTGAACCTTAAACG 657
Qy 221 LeuLeuLeuProIleTyrrAlaGlnAlaAlaAsnPheIleLeuAsnLeuLeuGlnGlnGly 240
Db 658 CTATTACTACTATTTATGCGCAAGCTGCTAATTTTCAATTTAAATTTATTAACAACAGT 717
Qy 241 AlaGluLeuAlaAspGluThrAsnAlaAspIleIleAspSerGlnIleGluProAsnAla 260
Db 718 GCTGAATGGCTGATGATGAGATGAGATCGAATATACATCTTCAACAATGAACTATATCT 777
Qy 261 GlyThrSerAspAspTyrrTyrrLeuLeuGlySLeuIleProLySerSerAsnTyrr 280
Db 778 GGAACATCGATGATCTATTTATTAACCTTTAAAGAAATATCTTAATATATGTAACAT 837
Qy 281 CysAlaAsnThrTyrrArgThrGlyLeuLyAsnLeuAspGluProAsnMetLySerPro 300
Db 838 TGTGCAATATACCTATAGAGAGAGACTAAATAAATTGCAACGAACTATATGAGATGG 897
Qy 301 SerIlePheAsnAspTyrrArgArgTyrrMetThrIleThrValLeuAspThrIleSerGln 320
Db 898 AATAATATTTAATGATTAATCGAAGATATATGACTATTTATCTATTTGATATCTATCCCTCA 957
Qy 321 PheSerLeuTyrrAspIleLySerGlyTyrrArgAspSerIleGlyGlyIleGluValIleGly 340
Db 958 TTTTCTTTTATGATTAATAGATACAAAGATTCATATGGA-----AGAAATAGTGGC 1011
Qy 341 IleLyAsnGluLeuThrArgGluIleTyrrThrGlnIleAsnPheAspArgLeuPro 360
Db 1012 ATTTAAACGGAACCTTACAGAGAAATTTATACACGAAATAAATTTTGGACCGTCTTACT 1071
Qy 361 GlnLeuArgValGlnProAsnLeuAlaThrMetGlyTyrrAsnLeuThrArgAlaSerPhe 380
Db 1072 TACCTTGAATTCACCCCAATCTCGCTATAATGAGATTAATTAATTAACGTTTCAGGGCTT 1131
Qy 381 LysIleuPheSerPheLeuGlnGlnPheIlePheTyrrThrGlnAsnThrAsnPheGlyAsn 400
Db 1132 ACATTAATTTTCAATTTTATGATGAACTTATATTTTATACAAAATGAAACGTAAGGGAAT 1191
Qy 401 ArgLeuValGlyIleSerAsnArgAspAlaProThrTyrrSerAsnThrIleThrGluThr 420

Db 1192 CGTTTAGTGGTATTCGCAATCGTAATGATCTACTATGCTACGACGAGAACTGAATT 1251
Qy 421 LeuTyrrGlyIleArgThrGlySerProThrThrTyrrThrIleArgProPheGluSerTyrr 440
Db 1252 ATATATGAGAAAGAACAGGTCCACCCACAAACAAATCTTAATATACATTTGAAATCTCAT 1311
Qy 441 LysValSerIleValThrAspArgGlnSerProProValSerProIle--GlnProHis 459
Db 1312 AAAGTTTCAATTTGAACTGATAGCAAGTAACCTCTTATCTCCCTTTTCTTAACATATAC 1371
Qy 460 PheIleIleAsnGlnIleGluLeuTyrrLeuAsnGlySerSerAsnAsnThrLeuLySerTyrr 479
Db 1372 TTACAAATTAATCAAAATTTGAACCTTTATTTAATTAATTCACCTAGTAATTAATTAACATAT 1431
Qy 480 SerAlaGlySerLeuSerAsnTyrrGlnAsnThrThrPhePheGlnPheProArgLyAs 499
Db 1432 TCAGCTGGGGGAATTTATCTATGATTAATAAAACAACTGATTTCAATTTCCGTATAAA 1491
Qy 500 LysAspCysAsnLeuValIleAspProGlyCysSerProAsnPheAsnAsnTyrrSerHis 519
Db 1492 AAAGCTGTAACCAATATTAATTCAAATTTGTTTACCAAGCTAATTAATGTTATATGTCAT 1551
Qy 520 IleLeuSerHisPheSerLeuPheThrTyrrSerTyrrValIleGlyLeuGlnLeuGlnIle 539
Db 1552 ATTTATCCCGATTTTCTTATTTAATTAATTTCTTAATAAATTTGATTAGCCGCTAAATATA 1611
Qy 540 LeuAspThrGlyValIleGlyTyrrThrHisSerSerValaAspArgTyrrAsnAlaIleSer 559
Db 1612 TTATATACAGGTGCATTTAGATGAGACACAGTATGTTATTAAGAAATTAATGCAATATCA 1671
Qy 560 AspLyIleIleIleThrMetIleProAlaIleLyGlyAsnAsnLeuAspThrAsnSerLyS 579
Db 1672 GATTAATAATTAATTAATGATATCCAGCATCAAGGTAAACAGCTCTTGAATCAAACTCTAAG 1731
Qy 580 ValIleGlyGlyProGlyHisIleThrGlyLyAsnLeuValTyrrLeuGlnSerGlnGlyArg 599
Db 1732 GTAATTTGAAGACCTGTCATACAGAGAGAACTTGTTTATTTTACAAAGTCACAGGCGCT 1791
Qy 600 LeuGlnIleThrCysGluThrProAsnSerThrGlnSerTyrrPheIleArgLeuArgTyrr 619
Db 1792 TTACAGATTAATGATGAACTCCATATTTCTACACATTTTATTAATTAATTAATTAATTAAT 1851
Qy 620 AlaThrAsnGlyAlaGlyAsnThrLeuProAsnIleSerLeuThrIleProGlyValIle 639
Db 1852 GCTACAAATGGTGCTGGAATATCTCTCTATATATCTTTCAACATCCAGAGATATA 1911
Qy 640 GlyIleProProGlnArgLeuAsnAsnThrPheSerGlyTyrrAsnTyrrAsnAsnLeuGln 659
Db 1912 GGAATACCACTCAACGACTCAACAAACCTTTTCTGTACAAATTAATTAATTAATTAATTAAT 1971
Qy 660 TyrrGlyAspPheGlyTyrrPheGlnPheProSerThrValThrLeuProLeuAsnArgAsn 679
Db 1972 TACGAGATTTTGGTATTTTCAATTTTCCAAAGTACAGTAACATTTACTTTAAATGGAAC 2031
Qy 680 IleProPheIlePheAsnArgAlaAspValSerAsnSerIleLeuIleIleAspLySerIle 699
Db 2032 ATACATTTATATTAATTAATCGGACAGATGATCAAAATTCATTTTAAATCATTTGATTAAT 2091
Qy 700 GluPheIleProIleThrSerSerMetHisGlnAsnArgGlyGlyGlnLyLeuGlnThr 719
Db 2092 GAATTTAATCAAAATTAATTAATCTCTCTGTCGCCAAAATTAAGAAAACAAAATTTAGAAACT 2151
Qy 720 IleGlnThrLyIleAsnThrPhePheThrAsnHisIleThrLyS 733
Db 2152 ATCCAAACAAAATAATTAATCAATTTTTCACAAATCATACAAA 2193
RESULT 3
ID ADR89398 standard; cDNA; 2085 BP.
AC ADR89398;
XX
DT 18-NOV-2004 (first entry)

XX AXMI-007 alternative start site coding sequence.
DE ss; gene; delta-endotoxin; delta-endotoxin associate polypeptide;
XX expression cassette; transformation; transgenic; plant; bacteria;
KM lepidoptera; coleoptera; pest; pesticide; resistance;
XX pesticidal activity.
OS *Bacillus thuringiensis*.
FH Key Location/Qualifiers
FT CDS 1..2085
FT /tag=a
FT /product="AXMI-007"
PN WO2004074462-A2.
XX 02-SEP-2004.
PD 20-FEB-2004; 2004WO-US005829.
XX 20-FEB-2003; 2003US-0448632P.
XX 20-FEB-2003; 2003US-0448633P.
XX 20-FEB-2003; 2003US-0448797P.
XX 20-FEB-2003; 2003US-0448806P.
XX 20-FEB-2003; 2003US-0448810P.
XX 20-FEB-2003; 2003US-0448812P.
XX 19-FEB-2004; 2004US-00781979.
XX 19-FEB-2004; 2004US-00782020.
XX 19-FEB-2004; 2004US-00782096.
XX 19-FEB-2004; 2004US-00782141.
XX 19-FEB-2004; 2004US-00782570.
XX 19-FEB-2004; 2004US-00783417.
XX (ATHE-) ATHENIX CORP.
XX Carozzi N, Hargies T, Koziel MG, Duck NB, Carr B;
PI WPI: 2004-635574/61.
XX P-PSDB; ADR89399.
XX New isolated delta-endotoxin and delta-endotoxin-associated nucleic acids
PT and polypeptides, useful for killing lepidopteran or coleopteran pests or
PT for producing organisms with pesticide resistance.
XX Claim 1; SEQ ID NO 10; 178bp; English.
XX This sequence encodes an isolated delta-endotoxin. Some of the delta-
CC endotoxin coding sequences of the invention have alternative start
CC codons, producing more than one protein from a single open reading frame.
CC The nucleic acid sequences of the invention are useful in DNA constructs
CC or expression cassettes for transformation and expression in plants and
CC bacteria. The nucleic acids and corresponding polypeptides are useful for
CC killing lepidopteran or coleopteran pests. Compositions containing the
CC delta-endotoxins of the invention, and methods for their production, are
CC useful for the production of organisms with pesticide resistance,
CC specifically bacteria and plants. These organisms are useful for
CC generating altered or improved delta-endotoxin or delta-endotoxin-
CC associated proteins that have pesticidal activity, or for detecting the
CC presence of delta-endotoxin or delta-endotoxin-associated proteins or
CC nucleic acids in products or organisms.
XX Sequence 2085 BP; 794 A; 338 C; 302 G; 651 T; 0 U; 0 Other;
SQ

Alignment Scores:
Pred. No.: 7.9e-271 Length: 2085
Score: 3104.50 Matches: 598
Percent Similarity: 91.9% Conservative: 30
Best Local Similarity: 87.6% Mismatches: 52
Query Match: 80.2% Indels: 3
DB: 13 Gaps: 2

US-10-783-417-2 (1-735) * ADR89398 (1-2085)

QY 52 MetCysGInGlyAenThrGlnTyrGlyAspAsnPheGluThrPheAlaSerAlaAspThr 71
DB 1 ATGTCTCAAGGGAATACCAATATGCTGATATTTCCAGACATTTGCTAGTCTGATACA 60
QY 72 TLeAlaAlaValSerAlaGlyThrIleValSerGlyThrLeuLeuAlaGlyIleGlyGly 91
DB 61 ATTGCTGAGTTAGTACGAGTACTATGATCCGGTACTCTGTATGCCGATATAGTGGG 120
QY 92 LeuThrSerIleSerGlyProIleGlyIleIleGlyAlaIleIleIleSerPheGlyThr 111
DB 121 CTCACCTTCAATACCGACCAATAGAAATATAGGTGCTATATATATATCTTTTGCTACC 180
QY 112 LeuIleThrValPheThrProAlaGlyGluGlnAspLysThrValIleThrGlnPheIle 131
DB 181 CTAATCAGCTGCTTTTGGCCGCGGAGAACACAAACAAAGATGACACAAATTTATT 240
QY 132 LysMetGlyGluIlePheValAspThrProLeuThrGlnSerIleGlyGlnLeuLysLeu 151
DB 241 AAAATGGAGAAATTTTGTGATACACCGTTAACAAAGATTAACAGCTAAAGTTA 300
QY 152 GlnThrLeuGluGlyPheArgGlnIleLeuGlnSerTyrAsnThrAlaLeuAspAspThr 171
DB 301 CAACCTTGAAGAGATTATGACAAATATTACAAAGCTATATACGACTTATGATGATTGG 360
QY 172 ArgLysLeuLysArgLeuGlnAlaProGlyLeuProPheSerAlaLeuGlnGlnAla 191
DB 361 AGAAATATTAAGAGCTACAGCTCTGATATTCACCACTATCAGCATTTACACAGAGCT 420
QY 192 AlaLeuThrLeuLysIleArgPheGluAsnValHisAsnAspPheIleArgGluIlePro 211
DB 421 GCCTTGACTCTTAAATACGANTTACAGATGTTCAATGANTTTTATTCGAAATATCCT 480
QY 212 GlyPheGlnLeuGluThrTyrLysThrLeuLeuLeuProIleTyrAlaGlnAlaAsn 231
DB 481 GGTTCCAACTTGAAACTTATTAACGCTATTACTTACTTATTAATGCGCAAGCTCTAAT 540
QY 232 PheHisLeuAsnLeuLeuGlnGlnGlyAlaGluLeuAlaAspGluTTPAsnAlaAspIle 251
DB 541 TTTCAATTAAATTTATTAACACAGAGCTGATATGCGCAATGAGAAATGCAATTA 600
QY 252 HisProSerGlnIleGluProAsnAlaGlyThrSerAspAspTyrTyrLysLeuLys 271
DB 601 CATCTTCACAAATTTGAACCTATCTGAGACATGAGTACTATTAATCTTTTAAA 660
QY 272 GluAsnIleProLysTyrSerAsnTyrCysAlaAsnThrTyrArgThrGlyLeuLysAsn 291
DB 661 GAAATATATACCTTAATATATAGTAACTATGTGCAATTAACCTATAGAGAAAGCACTAATAA 720
QY 292 LeuArgAspGluProAsnMetLysTrpSerIlePheAsnAspTyrArgTyrMetThr 311
DB 721 CTTCCAAACGAACTTAATATATGAGATATATTAATGATTATTCGAAGATATATGACT 780
QY 312 IleThrValLeuAspThrIleSerGlnPheSerLeuTyrAspIleLysArgTyrArgAsp 331
DB 781 ATTACTGATTTAGTACTATCTCAATTTCTTTTATGATATTAAGATACAAAGAT 840
QY 332 SerIleGlyGlyIleGluValIleLysValIleLysAsnGluLeuThrArgGluIleTyrThr 351
DB 841 TCAATATAGG-----AGATAGGTGTCATTAAACGAACTTAACAGAAATTTATACA 894
QY 352 ThrGluIleAsnPheAspArgLeuProGlnLeuAlaGlnProAsnLeuAlaThrMet 371
DB 895 ACTGAATTAATATTTGACCGTCTTACTTACTTGAATTAACCAACCAATCTCGCTATTAAG 954
QY 372 GluTyrAsnLeuThrArgAlaSerPheLysLeuPheSerPheLeuGlnPheIlePhe 391
DB 955 GAAATATATTTATACAGTTCAGGCGCTTAGATTTATTTTATTTTATAGTAACTTATATT 1014
QY 392 TyrThrGluAsnThrAsnPheGlyAsnArgLeuValGlyIleSerAsnArgAspAlaPro 411
DB 1015 TATACAAAAAATGAACGATACCGGAAATCGTTAGTGTGATTCGAAATCGTAAATGATCT 1074

```

OY 412 ThrTyrSerAsnThrIleThreGluThrLeuTyrGlyGluAspThrGlySerProThrThr 431
Db 1075 ACTTATGCTACGACGAACTGAAATATATATGAGAAAGAAAGGCTCCACCCCAACA 1134
OY 432 LysThrIleArgProPheGluSerTyrIleValSerIleValThrAspArgIleSerPro 451
Db 1135 AAAACCTTAAATCCATTGAACTCTTAAGTTTCAATTGTTACTGATAGACAGTAAGT 1194
OY 452 ProValSerProIle--GlnProHisPheIleIleAsnGlnIleGluLeuTyrLeuAsn 470
Db 1195 CCTACTCCCTCTTCTTCAACATATACCTTACAAATTAATCAATTAATGAATTTAATTAAT 1254
OY 471 GlySerSerAsnAsnThrLeuLysTyrSerIleArgIleGlySerLeuSerAsnTyrGlnAsn 490
Db 1255 AATTACCTAGTAATTAATTAATCAATTAATCAATTAATCAATTAATTAATTAATTAAT 1314
OY 491 ThrThrPhePheGlnPheProArgLysIleAspGlySerIleValIleAspProGlyCys 510
Db 1315 ACRACTGATTTTCAATTTCTCTTAATAAAGACGTTAAACCAATTATTAATCAAAATTTG 1374
OY 511 SerProAsnPheAsnAsnTyrSerHisIleLeuSerHisPheSerLeuPheThrTyrSer 530
Db 1375 TTACCAAGCTAATATAGTATATGTCATATTTATCCAGTTTCTTTATTTAATTAATTC 1434
OY 531 TyrValIleGlyLeuGlnLeuGlnIleLeuAspThrGlyValLeuGlyTyrThrHisSer 550
Db 1435 TATATAAATGATGTTAGCGCTAATATATATATATATACAGTGATGAGATGACACACAGT 1494
OY 551 SerValAspArgTyrAsnAlaIleSerAspLysIleIleThrMetIleProAlaIleLys 570
Db 1495 AGTGTATATAGAAATATAGCAATATCAATATATATATATATATATATATATATATAT 1554
OY 571 GlyAsnAsnLeuAspThrAsnSerLysValIleGluGlyProGlyHisThrGlyGlyAsn 590
Db 1555 GGTAACAGCTTGATACAACTCTAAGGTAATGAAAGACCGTGCATACAGGAAGAAAC 1614
OY 591 LeuValIleGlyLeuGlnSerGlnIleGlyLeuGlnIleThrCysGluThrProAsnSerThr 610
Db 1615 TTGGTTATTTTACAAAGTCAAAGGCGTTTACAGATTACATGTAGAACTCTTAATTTCTCA 1674
OY 611 GlnSerTyrPheIleArgLeuArgTyrAlaThrAsnGlyValAlaGlyAsnThrLeuProAsn 630
Db 1675 CAATCTTATACATTAGACTTGATGATACGCTACAAATGGTCTGGAATATCTCTTCATAT 1734
OY 631 IleSerLeuThrIleProGlyValIleGlyIleProProGlnArgLeuAsnAsnThrPhe 650
Db 1735 ATATCTTTACAAATCCAGAGTATAGGAATACCACTCAACAGCTCAACAGACTTTT 1794
OY 651 SerGlyThrAsnTyrAsnAsnLeuGlnTyrGlyAspPheGlyTyrPheGlnPheProSer 670
Db 1795 TCTGGTACAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1854
OY 671 ThrValThrLeuProLeuAsnArgAsnIleProPheIlePheAsnArgAlaAspValSer 690
Db 1855 ACAGTAACATTACCTTAAATGAAACATACCATTTATTTAATCGAGAGATGATCA 1914
OY 691 AsnSerIleLeuIleIleAspLysIleGluPheIleProIleThrSerSerMetHisGln 710
Db 1915 AATTCATTTTATATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1974
OY 711 AsnArgGlyLysGlnLysLeuGlnIleThrIleGlnThrLysIleAsnThrPhePheThrAsn 730
Db 1975 AATTAAGAAAAAATTAAGAACTATCCAAACAAATAATTAATTAATTAATTAATTAAT 2034
OY 731 HisThrLys 733
Db 2035 CATACAAA 2043

```

RESULT 4
 ID AAN93059 standard; DNA; 4571 BP.
 XX
 AC AAN93059;

```

XX 27-AUG-2003 (revised)
DT 25-MAR-2003 (revised)
DT 30-MAY-1990 (first entry)
XX
XX Delta-endotoxin crystal protein gene.
XX
XX Delta-endotoxin; crystal protein; insecticide; pCH130; ss;
XX biological control agent.
XX
XX Bacillus thuringiensis; israelensis.
XX
XX
XX Key Location/Qualifiers
XX RBS 879..884
XX /tag= b
XX /label= Shine-Dalgarno sequence.
XX misc_feature 891..4430
XX /*tag= a
XX /*product= "delta-endotoxin crystal protein"
XX
XX EP296870-A.
XX
XX
XX 28-DEC-1988.
XX
XX 24-JUN-1988; 88EP-00305772.
XX
XX 26-JUN-1987; 87US-00067653.
XX
XX (DUPO ) DU PONT DE MEMOURS & CO E. I.
XX
XX Ellar DJ, Ward ES;
XX
XX WPI; 1989-001322/01.
XX P-PSDB; AAP93715.
XX
XX DNA fragment encoding insecticidal protein - obtd. from Bacillus
XX thuringiensis sub species israelensis, and used in microorganisms and
XX plant cells.
XX
XX Disclosure; Fig 7; 26pp; English.
XX
XX PS
XX CC The sequence encodes the 130 kDa delta-endotoxin gene from B.
XX CC thuringiensis subsp. israelensis and is a 4.46 kb insert of pCH130. The
XX CC delta endotoxin protein is insecticidal and can be used to control insect
XX CC pests esp. mosquitos. See also AAN93054. (Updated on 25-MAR-2003 to
XX CC correct PA field.) (Updated on 27-AUG-2003 to correct OS field.)
XX
XX SQ Sequence 4571 BP; 1678 A; 685 C; 817 G; 1391 T; 0 U; 0 Other;
XX
XX Alignment Scores:
XX Pred. No.: 2.48e-85 Length: 4571
XX Score: 1061.50 Matches: 271
XX Percent Similarity: 50.4% Conservative: 118
XX Best Local Similarity: 35.1% Mismatches: 278
XX Query Match: 27.4% Indels: 105
XX DB: 1 Gaps: 24
XX
XX US-10-783-417-2 (1-735) x AAN93059 (1-4571)
OY 1 MetAsnGlnAsnAsnAspAsnAsnGluTyrGluIleIleAspSerHisThrSerProTyr 20
Db 891 ATGAAATCCTTATCAAAATTAATAAATGAATATCAAAATTAATTAATTAATTAATTA 950
OY 21 PheProAsnArgAsnSerAspAspSerArgTyrProTyrThrAsnAsnProAsnGlnPro 40
Db 951 -----AATATATCTAATTAATTAATTAATCAAGATATCAATAGAAATATGCAAAAACAAATTA 1004
OY 41 LeuGlnAsnThrAsnTyrLysGluTyrPheAsnMetCysGlnGlyAsnThrGlnTyrGly 60
Db 1005 TTACAAAGTACAAATTAATTAAGATTTGGCTCAATATGTGTCAACGAATCAGACGATGTGT 1064
OY 61 AspAsnPheGluThrPheAlaSerAlaAspThrIleAlaIleValSerAlaGlyThrIle 80

```

Db 1065 GGAGATTTTGAACCTTTATGTAAGT-----GCTGAACCTACGCTATACTATT 1115
 QY 81 ValSerGlyThrLeuLeuAlaGlyIleGlyGlyLeuThrSerIleSerGlyProIleGly 100
 Db 1116 GTAGTTGGGACCGTACTGACTGTTTGGGTTCCACAACA-----CCCTTAGGA 1163
 QY 101 IleIleGlyAlaIleIleIleSerPheGlyThrLeuIleThrValPheTrpProAlaGly 120
 Db 1164 CTT-----GCTTAATAGGTTTGGTACATTAAACAGTCTTTTCCAGCCCA 1214
 QY 121 GluGlnAspLeuThrValIleTrpGlnPheIleLeuMetGlyGluIlePheValAspThr 140
 Db 1215 GACCAATCTAACACA---TGAGTGACTTTATTAACAACAATCTAAATAATTTATTAATAA 1271
 QY 141 ProLeuThrGluSerIleLeuGlnLeuLeuGlnThrLeuGlnGlyPheArgGlnIle 160
 Db 1272 GAAATGCAATCAACATATATTAAGTAATGCTAATTAATAATTTTAAACAGTCTTAAATGTT 1331
 QY 161 LeuGlnSerThrAsnThrAlaLeuAspAspTrpArgGlyLeuLeuArgLeuGlnAlaPro 180
 Db 1332 ATCAGACACTTATCATATACCTTAACAATGGAG-----AATATCCA 1376
 QY 181 GlyLeuProSerSerSerAlaLeuGlnAlaAlaLeuThrLeuLeuIleArgPheGln 200
 Db 1377 AACCCACAATAATACAGATGAAGACACAATCCAGCTAGTTCATTACCAATTTCAA 1436
 QY 201 AsnValHisAsnAspPheIleArgGluIlePro-----GlyPheGlnLeuGlnThrThr 218
 Db 1437 AATGTCATTCAGAGCTGTAACTCTGTCTCTCTAATCTAGTATTCAGATTCATATCTAT 1496
 QY 219 LysThrLeuLeuLeuProIleThrAlaGlnAlaAlaAsnPheHisLeuAsnLeuGln 238
 Db 1497 AACATATCTAGTATATCTAGTATATGCAACAAGCAACTTACATCTGCTGATTAAT 1556
 QY 239 GlnGlyAlaGluLeuAlaAspGluTrpAsnAlaAspIleHisProSerGlnIleGluPro 258
 Db 1557 CAAGCCGTCAAATTTGAAGCGTATTTAAACCAATCGAACAATTCATATATTAGAGCCT 1616
 QY 259 AsnAlaGlyThrSerAspAspTrpTrpLeuLeuLeuGlnAsnIleProLysTrpSer 278
 Db 1617 ---TTGCCAACAAGCAATGATATATATTCAGATTCAGATTCAGATTCAGATTCAGATTC 1673
 QY 279 AsnTrpCysAlaAsnThrTrpArgThrGlyLeuLeuAsnLeuArgAspGluPro----- 296
 Db 1674 AATTATTTGCTAACATCTATTAATAAAGATTAATTAATTAATAAAGACGCTGATAGT 1733
 QY 297 -----AsnMetLysTrpSerIlePheAsnAspTrpArgTrpMetThrIle 312
 Db 1734 AATCTTGATGAAATATTAACCTGAGACACATACATACGATTCGACAAATAATGACTACT 1793
 QY 313 ThrValLeuAspThrIleSerGlnPheSerLeuTrpAspIleLeuArgTrpArgAspSer 332
 Db 1794 GCTGATTTGATGCTTGTTCACCTCTTCTAATTAATGATGATGATGATGATGATGATGAT 1844
 QY 333 IleGlyGlyIleGluValLysGlyIleLysAsnGlnLeuThrArgGluIleThrThr 352
 Db 1845 -----CCAAATAGGTGCTCAATCTGAACTTACCTGAGAAATTTATTCAGGTA 1889
 QY 353 GluIleAsnPheAspArgLeuProGlnLeuArgValGlnProAsnLeuAlaThrMetGln 372
 Db 1890 ---CTTAATCTTGAGAAAGCCCTTAATAATTAAT-----GACTTCAATATCAAGAG 1940
 QY 373 TrpAsnLeuThrArgAlaSerPheLysLeuPheSerPheLeuGlnGlnPheIlePheTrp 392
 Db 1941 GATTCACCTTACACGTAGACCG---CATTTATTTACTGCTGATTCCTTGAATTTTAT 1997
 QY 393 -----ThrGluAsnThrAsnPheGlyAsnArg----- 401
 Db 1998 GAAAAAGCGCAACCTACTCTAATTAATTTTTCACCAAGCAATTAATATGTTTCACTTAC 2057
 QY 402 ---LeuValGlyIleSerAsnArgAspAlaProThrTrpSerAsnThrIleThrGluThr 420
 Db 2058 AACCTTGATTAATATATATCCAAAAAATCTAGGTTTGGAAATACAAATGAATCTGATTA 2117

QY 421 LeuTrpGlyGluArgThrGlySerProThrThrLysThrIleArgProPheGluSerTrp 440
 Db 2118 TTA-----AATCTCTGTTTGGCAACAATATTATTAAT----- 2153
 QY 441 LysValSerIleValThrAspArgLysSerProValSerProIleGlnProHisPhe 460
 Db 2154 -----TTT 2156
 QY 461 IleIleAsnGlnIleGluLeu-----TrpLeuAsnGlySerSerAsnAsnThr--- 476
 Db 2157 TTATTAATATGTCATAGCTAGATTAATAATATCTAAATGATTAATTAATTAATTTGATA 2216
 QY 477 -----LeuLysTrpSerAlaGly 483
 Db 2217 ATGCAATTTTATATACATATAGTACTAGACTTTTGGAGAAAGAACTTACAGCAGATCT 2276
 QY 484 SerLeuSerAsnTrpGln---AsnThrThrPhePheGlnPheProArgLysAlaAspCys 502
 Db 2277 GCGCAATTAATCTATGATGATTAATAATAATATTTTCCGGTACCAATCTTAACGACAGA 2336
 QY 503 AsnLeuValIleAspProGlyCysSerProAsnPheAsnAsnTrpSerHisIleLeuSer 522
 Db 2337 GAGAACTCAAGAAACCTTACCTTTTCCACATATGATGATGATGATGATGATGATGATG 2396
 QY 523 HisPheSerLeuPheThrSerTrpValIleGlyLeuGlnLeuGlnIleLeuAspThr 542
 Db 2397 TTTATTAATAAGCTTATAGTATCCCTCCACATATTAATACTCAAGTAT----- 2444
 QY 543 GlyValLeuGlyTrpThrHisSerSerValAspArgTrpAsnAlaIleSerAspLysIle 562
 Db 2445 ---AGTTTGCCTGGACACACTCTAGTGTGATCCCTAAATAATTAATTAACACTTAA 2501
 QY 563 IleThrMetIleProAlaIleLysGlyAsnAsnLeuAspThrAsnSerLysValIleGlu 582
 Db 2502 ACTACCAAAATTCACGCTGTAAAGCGAATTCATCTGGAGCTCTTCAAGGTTGTCAA 2561
 QY 583 GlyProGlyHisThrGlyGlyAsnLeuValTrpLeuGlnSerGlnGlyArgLeuGluIle 602
 Db 2562 GGACCTGGTCATACAGAGGGGATTTAAT-----GATTCAAAGATCATTCAAAAT 2615
 QY 603 ThrCysGluThrProAsnSerThrGlnSerTrpPheIleArgLeuArgTrpAlaThrAsn 622
 Db 2616 ACATGTCAAACCTCAATTTTCAACATCGTATTTTATAGAAATTCGTATCTTCAAT 2675
 QY 623 GlyAlaGlyAsnThrLeuProAsnIleSerLeuThrIleProGlyValIleGlyIlePro 642
 Db 2676 GGAACGCGAAATATCTGAGCTGTATTAATCTTATGATCCAGGGGTACAGAACTG--- 2732
 QY 643 ProGlnArgLeuAsnAsnThrPheSerGlyThrAsnTrpAsnAsnLeuGlnTrpGlyAsp 662
 Db 2733 GGTATGGCACTAACCCCACTTTTCTGTGTACAGATTAACAAATTTAAATTAATAGAT 2792
 QY 663 PheGlyTrpPheGlnPheProSerThrValThrLeuProLeuAsnArgAsnIleProPhe 682
 Db 2793 TTTCACTAGTATGATTTTCTTACACGAGTGAATTTGCTCCAAATCAAAACATATCTCT 2852
 QY 683 IlePheAsnArgAlaAspVal---SerAsnSerIleLeuIleIleAspLysIleGluPhe 701
 Db 2853 GTGTTTAATCGTTCGATGTATATTAACAAACAAACAGTACTTATGATTAATTAATGAT 2912
 QY 702 IleProIleThrSerSerMetHisGlnAsnAsnArgGluLysGlnLysLeuGlnIleGln 721
 Db 2913 CTGCCAATTAATCTGTTCTATTAAGAGATTAAGAGAAACAAATAATTAAGAACAGTACAA 2972
 QY 722 ThrLysIleAsnThrPhePheThrAsnHisThrLys 733
 Db 2973 CAAATATTAATTAATCTTATATGCAATCTATTAATA 3008

RESULT 5
 AAQ14669
 ID AAQ14669 strand; DNA, 3543 BP.
 XX

AC AAQ14669;
 XX 27-AUG-2003 (revised)
 DT 25-MAR-2003 (revised)
 DT 04-FEB-1992 (first entry)
 XX
 DE Dipteran active toxin gene.
 XX
 KM Insecticide; B.t; crystal; delta endotoxin; cryIVA; ss.
 XX
 OS *Bacillus thuringiensis* serovar morrisoni.
 XX
 FH Key Location/Qualifiers
 FT CDS 1..3543
 FT /*tag= a
 XX
 PN EP457498-A.
 XX
 PD 21-NOV-1991.
 XX
 PF 09-MAY-1991; 91EP-00304180.
 XX
 PR 15-MAY-1990; 90US-00524255.
 PR 01-OCT-1990; 90US-00530903.
 XX
 PA (MYCO) MYCOGEN CORP.
 XX
 PI Sick AJ;
 XX
 DR WPI; 1991-341902/47.
 DR P-PSDB; AARI4373.
 XX
 PT *Bacillus thuringiensis* genes encoding dipteran-active toxins - and
 PT transformed microbes used to control insects in various environments.
 XX
 PS Claim 1; Page 10, 20pp; English.
 XX
 CC The sequence was obtd. from plasmid pMYC1625 which was isolated from a
 CC genomic library prepd. from DNA from B.t. PS71M3 [from B.t. PS71M3-69
 CC (NRRL B-1815)]. It is related to the cryIVA family of genes, the 140 kD
 CC endotoxin gene and the type II gene from B.t. var. israelensis. The gene
 CC encodes a 130 kD protein. Microorganisms transformed with the DNA may be
 CC administered to dipteran insects or their environments, the expressed
 CC toxins acting as an insecticide. See also AAQ14670-Q14672. (Updated on 25
 CC -MAR-2003 to correct PA field.) (Updated on 27-AUG-2003 to correct OS
 CC field.)
 CC
 SQ Sequence 3543 BP; 1278 A; 567 C; 612 G; 1086 T; 0 U; 0 Other;
 XX
 Alignment Scores:
 Pred. No.: 2.15e-85 Length: 3543
 Score: 1060.50 Matches: 271
 Percent Similarity: 50.34 Conservative: 117
 Best Local Similarity: 35.14 Mismatches: 279
 Query Match: 27.4% Indels: 105
 DB: 2 Gaps: 24
 US-10-783-417-2 (1-735) x AAQ14669 (1-3543)
 QY 1 MetAaNgInAaNsAaNsAaNgInUyYrGInUleIleAaSPSeRthIeRSeRProTyR 20
 DB 1 ATGAATCCCTTATCAAAATATAAATGAAATGAAACATTAATGCTTCACAAAAAAATTA 60
 QY 21 PheProAaNRgAaNsAaNsAaNsAaSPSeRArgTyRProTyRThRAaNsAaNsProAaNgInPro 40
 DB 61 -----AAATATATCTAATATATATATCAAGATATCCAAATAGAAAATAGTCCAAACCAATTA 114
 QY 41 LeuGInAaNsThRAaNsTyRLeGInUyRPlAaNsMeCySGInGlyAaNsThRgInTyRgLy 60
 DB 115 TTACAAAGACAAATATATTAAGATGGCTCAATATGTGCAACAGAAATGACAGATGTGT 174
 QY 61 AaPaAaNsPheGInUyRThRAaNsAaNsAaSerAlAaPThRIleAlaValSerAlaGlyThRIle 80
 DB ::::::::::::::::::::

DB 175 GGAGATTGTGAACCTTTATTGATAGT-----GGTGAACCTGACGCTTATACATAT 225
 QY ValSerGlyThRLeuLeuAlaGlyIleGlyLeuThRSeRileSerGlyProIleGly 100
 DB 81 :::::::::::::::::::::
 DB 226 GTAGTTGGACCGATCTACTGGTTGGGTTCCACACAC-----CCCTTAGA 273
 QY 101 IleIleGlyAlaIleIleIleSerPheGlyThRLeuIleThRValPheTrProAlaGly 120
 DB 274 CTT-----GCTTATATAGTTTGGTACATTAATACAGTCTTTTCCAGCCCA 324
 QY 121 GluGInAaNsPlySThRValTyRThRGlInPheIleIleIleSMeGlyGInUlePheValAaSPThR 140
 DB 325 GACCAATCTAACACA---TGAAGTACTTATTAACACAACTAAATAATTTATTAATAA 381
 QY 141 ProLeuThRGluseRileIleGlyLeuIleIleIleIleIleIleIleIleIleIleIleIle 160
 DB 382 GAATAGATCAACATATATATAGTAATCTAATATAAATTTTAAACAGGTCCGTTAATGTT 441
 QY 161 LeuGInSeRtyRAaNsThRAleuAaSPThRgLySleuIleAaRgLeuGInAlaPro 180
 DB 442 ATCAGCACTTATCAATATCACTTAAACATGGAG-----AATATCCA 486
 QY 181 GlyLeuProSeRSeRAlaLeuGInGInAlaAlaLeuThRLeuIleAaRgPheGly 200
 DB 487 AACCCACAAATPACTCAGATGTAGACACAAATCCAGTATTCATTAACATTTCAA 546
 QY 201 AaNsValHisAaAaSPheIleAaRgGlyIlePro-----GlyPheGInLeuGlyThRtyr 218
 DB 547 AATGTCATCCAGAGCTGTAACTCTGTCTCTCTTAATCCAGATGGATGGATTAACAT 606
 QY 219 LysThRLeuLeuProIleTyRAlaGInAlaAlaAaNsPheHileuAaNsLeuGIn 238
 DB 607 AACATACAGATATATCTAGTTATGACACACAGCAACCTTACATCTGATATTAAT 666
 QY 239 GInGlyAlaGlyLeuAlaAaSPGlyThRAaNsAlaAaSPleIleIleProSeRGlInIleGlyPro 258
 DB 667 CAAGCCGTCAAAATTTGAAGCGTATTTAAAAAACAATCCAGCAATTCGATTATTAAAGCCT 726
 QY 259 AaNsAlaGlyThRSeRAaSPtyrTyRtySleuLeuIleGlyAaNsIleProIletyrSeR 278
 DB 727 ---TTGCCAAGCAGCAATGATTTATTTATCCAGATTTAGCTTAAGATTAAGATTAACACT 783
 QY 279 AaNsTyRQyAlaAaNsThTyRAaRgThRgLyIleuIleAaNsAaRgAaSPGlyPro----- 296
 DB 784 AATTATGTGTAAACAATTAATAAAGATTAATTAATTAATAAACAACCGCTGATAGT 843
 QY 297 -----AaNsMeIleTyPSeRilePheAaNsAaSPtyrAaRgTyRThRThRIle 312
 DB 844 AATCTTGATGAAATATATTAACCTGGAACACATACATTAAGTATGAAACAAATGACTACT 903
 QY 313 ThRValLeuAaPThRIleSeRGlInPheSeRLeuTyRAaPileIleAaRgTyRgAaSPSeR 332
 DB 904 GCTGATTAATGATCTTTGGCACTCTTCTTAATTAATGATTAAGTAAATAT----- 954
 QY 333 IleGlyGlyIleGlyValIleGlyAlaIleIleAaNsGInUleuThRAaRgGlyIleTyRThR 352
 DB 955 -----CAATAGGTGTCCAATCTGAATCTTACGAGAAATTTATCAAGTA 999
 QY 353 GluIleAaNsPheAaRgLeuProGInUleuAaRgValGInProAaNsLeuAlaThRMeGly 372
 DB 1000 ---CTTAATCTCGAAGAAAGCCCTATTAATATAT-----GACTTCAATATCAAG 1050
 QY 373 TyRAaNsLeuThRAaRgAlaSeRPhelAaNsPheSeRPhelGluGInPheIlePheTyR 392
 DB 1051 GATTACTTACACGTAGACG---CATTTATTTATCTGGCTTGATTTCTTGAATTTTAT 1107
 QY 393 -----ThRGlusAnThRAaNsPheGlyAaNsAaRg----- 401
 DB 1108 GAAAAAGGCAACATACCTCTATATATTTTTCACACAGCATATATATATGTTTCATTAC 1167
 QY 402 ---LeuValAlaGlyIleSeRAaNsAaRgAaAlaProThTyRSeRAaNsThRIleThRGlUyR 420
 DB 1168 ACATTTGATATATATATCCAAATAATCTAGTTTGGAAATACAAATGATGATGATTA 1227

```

QY 421 LeuTYrGLyGLuArgThrGLySerProThrThyLstHrLeaArgProPheGluSerTYr 440
DB 1228 TTA-----AAATCTCTGGTGGCAACAAATATTATATT----- 1263
QY 441 LysValSerLeValThrAspArgInSerProValSerProIleGlnProHisPhe 460
DB 1264 -----TTT 1266
QY 461 IleIleGlnGlnIleGluLeu-----TYrLeuGlnGlySerSerAsnAsnThr--- 476
DB 1267 TTTATTAATAGTCATAGCTTAGATATATAATATCTAAATGATTAATATATTATTGATAA 1326
QY 477 -----LeuLYrTYrSerIleGlyGly 483
DB 1327 ATGCAATTTTATATACTAATAGTACTAGACTTTTGGAGAAAGACTTACAGACGAGATCT 1386
QY 484 SerLeuSerAsnTYrGln---AsnThrThrPhePheGlnPheProArgLysLysAspCys 502
DB 1387 GGGCAAAATACCTTAGATGATGATTAATAATAATTTTCGGTTTCGCAATTTCTTAACCAAGA 1446
QY 503 AsnLeuValIleAsnProGlyCysSerProAsnPheAsnAsnTYrSerHisIleLeuSer 522
DB 1447 GAGAAATCAAGCAATCCCTACCTTTTCCACATATGATATACATATATGATATTTATCA 1506
QY 523 HisPheSerLeuPheThrTYrSerTYrValIleGlyLeuGlnLeuGlnIleLeuAspThr 542
DB 1507 TTTATTTAAAGTCTTAGATCCCTGCAACATATATAAACTCAAGTGTAT----- 1554
QY 543 GlyValLeuGlyTYrThrHisSerSerValAspArgTYrAsnAlaIleSerAspLysIle 562
DB 1555 ---ACGTTTGGCTGGACACACTCTAGTGTGATCCGTAATAATACATTTATACACTTAA 1611
QY 563 IleHrMetIleProAlaIleLysGlyAsnAsnLeuAspThrAsnSerLysValIleGln 582
DB 1612 ACTACCAAAATTCAGCTGTAAAGGCAATTCACCTGGAGCTTCTTAAGCTTGTTCA 1671
QY 583 GlyProGlyHisThrGlyGlyAsnLeuValTYrLeuGlnSerGlnGlyArgLeuGlnIle 602
DB 1672 GGACCTGGTCATACAGAGGAGGATTTAAT-----GATTTCAAGATCATTTCAAAAT 1725
QY 603 ThrCysGluThrProAsnSerThrGlnSerTYrPheIleArgLeuArgTYrAlaThrAsn 622
DB 1726 ACATGTCAACACTCAAAATTTTCAACATGTATTTTATAGAAATTCGTTATGCTTCAAT 1785
QY 623 GlyAlaGlyAsnThrLeuProAsnIleSerLeuThrIleProGlyValIleGlyIlePro 642
DB 1786 GGAAGCGCAATACACAGAGCTGTATTAATCTTAGATCCAGGGGTACAGAACTG--- 1842
QY 643 ProGlnArgLeuAsnAsnThrPheSerGlyThrAsnTYrAsnAsnLeuGlnTYrGlyAsp 662
DB 1843 GGTATGGCACTCAACCCCACTTTTCTGCTACAGATTAACAAATTTAAATTAAGAT 1902
QY 663 PheGlyTYrPheGlnPheProSerThrValThrLeuProLeuAsnArgAsnIleProPhe 682
DB 1903 TTTCACTACTTAGAAATTTCTAACAGAGTAAATTTGCTTCAAAATCAAAACATCTCTT 1962
QY 683 IlePheAsnArgAlaAspVal---SerAsnSerIleLeuIleIleAspLysIleGluPhe 701
DB 1963 GGTATTAAATCGTTCCGATGTATATACAAACACAGAGTACTTATGATTAATTTGAAT 2022
QY 702 IleProIleThrSerSerMetHisGlnAsnArgGlyLysGlnLysLeuGlnThrIleGln 721
DB 2023 CTGCCAAATTAATCTGTTCTATAGAGAGTAGAGAGAAACAAAAATTTAGAAAACAGTACA 2082
QY 722 ThrLysIleAsnThrPhePheThrAsnHisThrLys 733
DB 2083 CAATATTAATTAATCATTTATGCAAAATCTATATAA 2118

```

RESULT 6
AA081178
ID AA081178 standard; DNA; 3543 BP.
XX

```

AC AA081178;
XX
DT 25-MAR-2003 (revised)
DT 12-AUG-1995 (first entry)
XX
DE B.t. toxin PS71M3 gene.
XX
KW Delta-endotoxin; crystal protein; biological control agent; Calliphorida;
KW screw-worm; sheep blowfly; Lucilia; Phormia; Calliphora; insecticide;
KW pesticide; B.t. 86.
XX
OS Bacillus thuringiensis.
XX
PN W09502694-A2.
XX
PD 26-JAN-1995.
XX
PF 13-JUL-1994; 94WO-US007902.
XX
PR 15-JUL-1993; 93US-00093199.
XX
PA (MYCO ) MYCOGEN CORP.
XX
PI Hickie LA, Payne J;
XX
DR MPI: 1995-067338/09.
DR P-PSDB; AAR63078.
XX
PT Method for controlling Calliphoridae pests - specifically utilises
PT Bacillus thuringiensis isolates or toxins.
XX
PS Disclosure, Page 36-38; 50pp; English.
XX
CC A library was constructed from Bacillus thuringiensis PS71M3 total
CC cellular DNA in lambda Gem-11. Plasmid pMYC1625, selected in Escherichia
CC coli, contained a 8.0 kb insert expressing a beta-endotoxin gene. This
CC was sequenced (AA081178). A cured, acrycallyiferous B.t. host carrying
CC pMYC1625 produced a 130 kDa crystal protein (AAR63078) toxic to
CC calliphorids. (Updated on 25-MAR-2003 to correct PN field.)
XX
SQ Sequence 3543 BP; 1278 A; 566 C; 612 G; 1087 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 2,15e-85 Length: 3543
Score: 1060.50 Matches: 271
Percent Similarity: 50.34 Conservative: 117
Best Local Similarity: 35.14 Mismatches: 279
Query Match: 27.44 Indels: 105
DB: Gaps: 24

US-10-783-417-2 (1-735) x AA081178 (1-3543)
QY 1 MetAngInAsnAsnAspAsnAsnGluTYrGluIleIleAspSerHisThrSerProTYr 20
DB 1 ATGAATCTCTTATCAAAATTAATAATATGAATTAACATTAAGCTTCACAAAAAAATTA 60
QY 21 PheProAsnArgAsnSerAsnAspSerArgTYrProTYrThrAsnAsnProAsnGlnPro 40
DB 61 -----AATATATCTAATTAATTAATACAGATATCAATATGAAATATGATCCAAACATTA 114
QY 41 LeuGlnAsnThrAsnTYrLysGluTYrLeuAsnMetCysGlnGlyAsnThrGlnTYrGly 60
DB 115 TTAACAAGTACAAATTAATAAGATTGGCTCAATATGTGCAACAGAAATCAGCATGTGCT 174
QY 61 AsnAsnPheGluThrPheAlaSerAlaAspThrIleAlaValSerAlaGlyThrIle 80
DB 175 GGAAGATTTTGAACCTTTATGATAGT-----GGTAACTCAGTGCCTATATCTATT 225
QY 81 ValSerGlyThrLeuLeuAlaGlyTYrIleGlyLysLeuThrSerIleSerGlyProIleGly 100
DB 226 GTAGTTGGGACCGGTACTGACTGTTTCCGGTTCCACAAAC-----CCCTTAGGA 273
QY 101 IleIleGlyAlaIleIleIleSerPheGlyThrLeuIleThrValPheThrProAlaGly 120

```



```

Db      274 CTT-----GCTTAAATAGTGTGGTACATTAATAACGACTCTTTTCCACCCCAA 324
Qy      121 GluGlnAspLysThrValTTrpThrGlnPheIleuSmetGlyLulLePheValAspThr 140
Db      325 GACCAATCTAACACA---TGGAGTGACTTTATACACAACTAAATAATTTATATAAAAAA 381
Qy      141 ProLeuThrGlnSerIleuSglnLeuLysLeuGlnThrLeuGlnGlyPheArgGlnIle 160
Db      382 GAATATGAGTCAACATATATATAAGTAATGCTATATAAATTTTAAACAGGTCGTTAATGTT 441
Qy      161 LeuGlnSerTyrAsnThrAlaLeuAspAspTTrpArgLysLeuLysArgLeuGlnAlaPro 180
Db      442 ATCAGACACTTATCATATATACACCTTAAACATGGGAG-----AATATATCA 486
Qy      181 GlyLeuProProSerSerAlaLeuGlnGlnAlaAlaLeuThrLeuLysIleArgPheGln 200
Db      487 AACCCACAAATATCTCAGAGATGTAAAGACACAAATCCAGCTAGTTCATTACCATTTTCA 546
Qy      201 AsnValHisAsnAspPheIleArgGluIlePro-----GlyPheGlnLeuGlnThrTyr 218
Db      547 AATGTCATTTCCAGAGCTTGTAACCTTGCTCTCTCTCTAATCCATGATTTGGCATTAACAT 606
Qy      219 LysThrLeuLeuLeuProIleTyrAlaGlnAlaAlaAsnPheHisLeuAsnLeuGln 238
Db      607 AACATACTAGTATATCTAGTATATCACAAGACGACAACTTACATCTGATCTAATAAT 666
Qy      239 GlnGlyAlaGlnLeuAlaAspGluTrpAsnAlaAspIleHisProSerGlnIleGluPro 258
Db      667 CAAAGCCGTCAAATTTGAAAGCGTATTTAAAAACAAATCCACAAATTCGATTTATTAAGCT 726
Qy      259 AsnAlaGlyThrSerAspAspTyrTyrLysLeuLeuLysGluAsnIleProLysTyrSer 278
Db      727 ---TTGCCAACAGCATTTGATTATATTCAGATTTAGACTAAAGCTATGAAGATTACACT 783
Qy      279 AsnTyrCysAlaAsnThrTyrArgThrGlyLeuLysAsnLeuArgAspGluPro----- 296
Db      784 AATATATGTGTAACTATATAAAGATTTAAATTTAATTTAAACGACGCTGATAGT 843
Qy      297 -----AsnMetLysTrpSerIlePheAsnAspTyrArgArgTyrMetThrIle 312
Db      844 AATCTTGATGGAATATATACTGGAACACATACAAATCGTATCGAACAAATATGACTCT 903
Qy      313 ThrValLeuAspThrIleSerGlnPheSerLeuTyrAspIleLysArgTyrArgAspSer 332
Db      904 GCTGATTAATGATCTGTTCGACTCTTCTCAATATATGATGATGATTAAT----- 954
Qy      333 IleGlyGlyIleGluValLysGlyLysLysAsnGlnLeuThrArgGlnIleTyrThrThr 352
Db      955 -----CCATAGGTGTCCAAATCTGAATCTTACTCGAATAATTATCAGCTA 999
Qy      353 GluIleAsnPheAspArgLeuProGlnLeuArgValGlnProAsnLeuAlaThrMetGlu 372
Db      1000 ---CTTAATCTTGAGAAAGCCCTATATAATTTAT-----GACTTCCTCAATATCAAG 1050
Qy      373 TyrAsnLeuThrArgAlaSerPheLysLeuPheSerPheLeuGlnGlnPheIlePheTyr 392
Db      1051 GATTCACCTTACCGTAGACG---CATTTATTTACTGGCTTGATCTTTGAATTTTAT 1107
Qy      393 -----ThrGluAsnThrAsnPheGlyAsnArg----- 401
Db      1108 GAAAAAGCCAAACTACTCTTAATTTTTCACACGCACTTATATATATGTTTCATTAC 1167
Qy      402 ---LeuValGlyIleSerAsnArgAspAlaProThrTyrSerAsnThrIlePheGlnThr 420
Db      1168 AACCTTGATATATATATCCAAATAATCTAGTGTGTTTGGAAATCACAATGATGATGATAA 1227
Qy      421 LeuTyrGlyGluArgThrGlySerProThrThrIleArgProPheGlnSerTyr 440
Db      1228 TTA-----AATCTCTGTGTTGGCAACAAATATTTATTT----- 1263
Qy      441 LysValSerIleValThrAspArgLysSerProValSerProIleGlnProHisPhe 460

```

```

Db      1264 -----TTT 1266
Qy      461 IleIleAsnGlnIleGluLeu-----TyrLeuAsnGlySerSerAsnAsnThr--- 476
Db      1267 TTATTAATATGTCATTAAGCTTATGATTAATATATCTAATATGATTAATATATTTAGTAAA 1326
Qy      477 -----LeuLysTyrSerAlaGly 483
Db      1327 ATGATTTTATTAATACTAATGTACTAGTACTTTTGAGAAAGAACTTACAGCAGATCT 1386
Qy      484 SerLeuSerAsnTyrGln---AsnThrThrPhePheGlnPheProArgLysLysAspCys 502
Db      1387 GGGCAATTAACCTTATGATGTAATATAAATAATTTTGGGTTACCAATTTCTTAACCAAGA 1446
Qy      503 AsnLeuValIleAspProGlyCysSerProAsnPheAsnAsnTyrSerHisIleLeuSer 522
Db      1447 GAGATTCAGCAATCCCTACCTTTTCCACATATGTATACATATGATATGATTTATATCA 1506
Qy      523 HisPheSerLeuPheThrTyrSerTyrValIleGlyLeuGlnIleLeuAspThr 542
Db      1507 TTTATTAATAAGTCTTAGTATCCCTGCACATATATAAACTCAAGTAT----- 1554
Qy      543 GlyValLeuGlyTTrpThrHisSerSerValAspArgTyrAsnAlaIleSerAspLysIle 562
Db      1555 ---ACGTTTGCTTGACACACTCTAGTGTGATCTTAAATAATCAATTTATACACATTTA 1611
Qy      563 IleThrMetIleProAlaIleLysGlyLysAsnLeuAspThrAsnSerLysValIleGlu 582
Db      1612 ACTACCCAAATTTCCAGCTGTATTAACGCAATTCATCTGGAGACTGCTCTTAAGGTTCTCA 1671
Qy      583 GlyProGlyHisThrGlyLysAsnLeuValTyrLeuGlnSerGlnLysArgLeuGluIle 602
Db      1672 GACCTGGTCAATACAGAGGGGATTTAAT-----GATTCAAAGATCATTTCAAAAT 1725
Qy      603 ThrCysGluThrProAsnSerThrGlnSerTyrPheIleArgLeuArgTyrAlaThrAsn 622
Db      1726 ACATGTCAACACTCAAAATTTTCAACATCGTATTTATTAATAATTCGTTATGCTTCAAT 1785
Qy      623 GlyAlaGlyAsnThrLeuProAsnIleSerLeuThrIleProGlyValIleGlyIlePro 642
Db      1786 GAAAGCGCAATATACAGAGCTGTATTAATCTTACTATCCAGGGGTAGCAGAACTG--- 1842
Qy      643 ProGlnArgLeuAsnAsnThrPheSerGlyThrAsnTyrAsnAsnLeuGlnTyrGlyAsp 662
Db      1843 GGTATGAGCATCAACCCACTTTTCTGTGATACATTTATAGAAATTTAAATATATAAGAT 1902
Qy      663 PheGlyTyrPheGlnPheProSerThrValThrLeuProLeuAsnArgAsnIleProPhe 682
Db      1903 TTTCACTACTTATAGAAATTTTCTTAACGAGGTGAATTTGCTCCAAATCAAAACATATCTCT 1962
Qy      683 IlePheAsnArgAlaAspVal---SerAsnSerIleLeuIleIleAspLysIleGluPhe 701
Db      1963 GGTGTTAATCGTTCGATGTATATACAAACACACAGACTTATTTGATTAATAATGAAATTT 2022
Qy      702 IleProIleThrSerSerMetHisGlnAsnArgGlyLysGlnLysLeuGlnThrIleGln 721
Db      2023 CTGCAATTAATCTCGTTCTATAGAAGAGATAGAGGAACAAATTTGAAACAGTACAA 2082
Qy      722 ThrLysIleAsnThrPhePheThrAsnHisThrLys 733
Db      2083 CAAATTAATTAATACATTTATGCAAAATCTTATAAAA 2118

```

RESULT 7
 AAN93054
 ID AAN93054 standard, DNA, 3940 BP.
 AC AAN93054;
 AC XX
 DT 27-AUG-2003 (revised)
 DT 25-MAY-2003 (revised)
 DT 30-MAY-1990 (first entry)
 XX
 XX Delta-endotoxin crystal protein gene.

XX Delta-endotoxin; crystal protein; insecticide; ss; PCC130;
KM biological control agent.
XX
OS Bacillus thuringiensis; israelensis.
XX
FH Key Location/Qualifiers
FT RBS 879. .884 /*tag= b
FT /*label= Shine-Delgarno sequence
FT CDS 891 /*tag= a
FT /*product= "delta-endotoxin"
PN BP296870-A.
XX
PD 28-DEC-1988.
XX
PF 24-JUN-1988; 88EP-00305772.
XX
PR 26-JUN-1987; 87US-00067653.
XX
PA (DUPO) DU PONT DE NEMOURS & CO E I.
PI Ellar DJ, Ward ES;
PI
XX WPI: 1989-001332/01.
DR P-PSDB; AAP94035.
XX
XX DNA fragment encoding insecticidal protein - obcd. from Bacillus
PT thuringiensis sub species israelensis, and used in microorganisms and
PI plant cells.
XX
PS Disclosure; Fig 5; 26pp; English.
XX
CC The nucleotide sequence is an insert in plasmid PCC130. The delta
CC endotoxin protein is insecticidal and can be used to control insect pests
CC esp. mosquitoes. See also AAN93059. (Updated on 25-MAR-2003 to correct PA
CC field.) (Updated on 27-ANG-2003 to correct OS field.)
SQ Sequence 3940 BP; 1457 A; 603 C; 659 G; 1221 T; 0 U; 0 Other;
XX
XX
Alignment Scores:
Pred. No.: 3.06e-85 Length: 3940
Score: 1059.50 Matches: 271
Percent Similarity: 50.3% Conservative: 117
Best Local Similarity: 35.1% Mismatches: 279
Query Match: 27.4% Indels: 105
DB: 1 Gaps: 24
US-10-783-417-2 (1-735) x AAN93054 (1-3940)
QY 1 MetAsnGlnAsnAsnAsnAsnAsnGluTyrGluIleIleAspSerHisThrSerProTyr 20
DB 891 ATGAATCCCTTAATCAAAATTAATAATGAATATGAATCACTTAATGCTTCACAAAATAAATTA 950
QY 21 PheProAsnArgAsnSerAsnAspSerArgTyrProTyrThrAsnAsnProAsnGlnPro 40
DB 951 -----AATATATCTAATAATATTATACAGATATCCATAGAAATAGTCCAAAACAATTA 1004
QY 41 LeuGlnAsnThrAsnTyrLeuGluTyrLeuAsnMetCysGlnGlyAsnThrGlnTyrGly 60
DB 1005 TTACAAAGTACAAATTAATAAGATGCTCAATATATGTCACAGATCAGCGATAGGT 1064
QY 61 AspAsnPheGluThrPheAlaSerAlaAspThrIleAlaAlaValSerAlaGlyThrIle 80
DB 1065 CGAGATTTGAAACTTTTATATGATAGT-----GGTGAAGTCAAGTGCCTATATCTATT 1115
QY 81 ValSerGlyThrLeuLeuAlaGlyIleGlyGlyLeuThrSerIleSerGlyProIleGly 100
DB 1116 GTAGTTGGGAGCCCTACTGACTGGTTCCGGGTTCCACACA-----CCCTTAGGA 1163
QY 101 IleIleGlyAlaIleIleIleSerPheGlyThrLeuIleThrValPheTrpProAlaGly 120

DB 1164 CTT-----GCTTAATAGGTTTGGTACATTATATACAGTTCTTTTCAGCCCAA 1214
QY 121 GluGlnAspGlyThrValThrPheGlnPheIleYsMetGlyGluIlePheValAspThr 140
DB 1215 GACCAATCTMACACA---TGAGTGCCTTATTAACAACAATAATATTATTAACAAAA 1271
QY 141 ProLeuThrGluSerIleYsGlnLeuYsLeuGlnThrLeuGlnGlyPheArgGlnIle 160
DB 1272 GAAATRGCATCAACATATATTAAGTAATGCTAATTAATTTAAACAGGTCCTTAATGTT 1331
QY 161 LeuGlnSerTyrAsnThrAlaLeuAspAspThrPArgGlyLeuYsArgLeuGlnAlaPro 180
DB 1332 ATCAGACTTATCATATATACCTTAACATGGAG-----AATATCCA 1376
QY 181 GlyLeuProProSerSerAlaLeuGlnGlnAlaAlaLeuThrLeuYsIleArgPheGlu 200
DB 1377 AACCCACAAAATATCAGAGATGTAAAGACACAATTCAGCTAGTTCATTACCATTTTCAA 1436
QY 201 AsnValHisAsnAspPheIleArgGluIlePro-----GlyPheGlnLeuGluThrTyr 218
DB 1437 AATGTCATTCCACAGCTTGTAACCTTGCTCTCTAATCTAGTGAATTCGATTAATCTAT 1496
QY 219 LysThrLeuLeuLeuProIleTyrAlaGlnAlaAlaAsnPheHisLeuAsnLeuLeuGln 238
DB 1497 AACATATCTAGTATTTATCTAGTATATGACACAGACGAACTTACATCTGACTATTAAT 1556
QY 239 GlnGlyAlaGluLeuAlaAspGluTyrPheAlaAspIleHisProSerGlnIleGluPro 258
DB 1557 CAGCCGTCGTAATTTGAACGCTATTTAAACAAATCGAACATTCGATTATTAAGAGCCT 1616
QY 259 AsnAlaGlyThrSerAspAspTyrTyrIleLeuLeuYsGluAsnIleProIleTyrSer 278
DB 1617 ---TTGCCAACAGCAATTGATTATTAATTCAGATTTGACATTAAGACTTAAGATTACACT 1673
QY 279 AsnTyrCysAlaAsnThrTyrArgThrGlyLeuYsAsnLeuArgAspGluPro----- 296
DB 1674 AATTATGTGTAAACAACATTATTAATAAAGATTAATTTAATTAACAGCGCTGATAGT 1733
QY 297 -----AsnMetLysTrpSerIlePheAsnAspTyrArgArgTyrMetThrIle 312
DB 1734 AATCTTGATGAAATTAATAAATCGGAACATCAATCAATACGTAACAAATAATGACTACT 1793
QY 313 ThrValLeuAspThrIleSerGlnPheSerLeuTyrAspIleLeuYsArgTyrArgAspSer 332
DB 1794 GCTGATTAAGATGTTGTGACCTCTTCTTAATTTATGATGATGATTAATAT----- 1844
QY 333 IleGlyGlyIleGluValYsLeuYsAsnGluLeuThrArgGluIleTyrThrThr 352
DB 1845 -----CCAAATAGGTGTCCAATCTGAATCTGAATCTGAGAAATTTATTCAGGTA 1889
QY 353 GluIleAsnPheAspArgLeuProGlnLeuArgValGlnProAsnLeuAlaThrMetGlu 372
DB 1890 ---CTTAATCTTGAAAGAAAGCCCTTAATAATTTAT-----GACTTCAATATCAAGAG 1940
QY 373 TyrAsnLeuThrArgAlaSerPheYsLeuPheSerPheLeuGlnPheIlePheTyr 392
DB 1941 GATTCATCTTACAGTGAACCG---CATTTATTTACTGCTGATTCCTTGAATTTTAT 1997
QY 393 -----ThrGluAsnThrAsnPheGlyAsnArg----- 401
DB 1998 GAAAAAGCCGAACACTACTCTTAATATTTTTCACACGCAATTAATATGTTTCACTTAC 2057
QY 402 ---LeuValGlyIleSerAsnArgAspAlaProThrTyrSerAsnThrIleThrGluThr 420
DB 2058 ACACCTGATTAATATATCCCAAAATCTAGTGTTTTGGAAACACAACTAATCTAGTATAA 2117
QY 421 LeuTyrGlyGluArgGlySerProThrTyrIleThrIleArgProPheGluSerTyr 440
DB 2118 TTA-----AATCTCTTGTTTGGCAAAATATTTATATTT----- 2153
QY 441 LysValSerIleValThrAspArgGlnSerProValSerProIleGlnProHisPhe 460

```

Db 2154 -----TTC 2156
QY 461 ILeIleAsnGlnIleGluLeu-----TyrLeuAsnGlySerSerAsnAsnThr--- 476
XX :::::|||||
XX 2157 TTATTAAATGTCATAGCTTATAGATATAATAATCTTAATGATTATATATAGTAA 2216
QY 477 -----LeuLysTyrSerLysIleGly 483
XX :::::|||||
Db 2217 ATGATTTTTTTTAACTAATAGTACTAGACTTTTGGAGAAAAGAACTTACACGAGATCT 2276
QY 484 SerLeuSerAsnTyrGln---AsnThrThrPhePheGlnPheProArgLysLysAspCys 502
XX :::::|||||
Db 2277 GGGCAAAATTAAGTATGATTAATAATAATATTTCCGGTTTCCAAATCTTAAACGACAGA 2336
QY 503 AsnLeuValIleAspProGlyCysSerProAsnPheAsnAsnTyrSerHisIleLeuSer 522
XX :::::|||||
Db 2337 GAGAAATCAAGAAACCTTACCTTTTCCAAACATATGATTAATAGTATATTTTATCA 2396
QY 523 HisPheSerLeuPheThrTyrSerTyrValIleGlyLeuGlnIleLeuAspThr 542
XX :::::|||||
Db 2397 TTTATTAAAGTCTTATGATCCCTGCACATATAAACTCAAGTGTAT----- 2444
QY 543 GlyValIleGlyTyrThrHisSerSerValAspArgTyrAsnAlaIleSerAspLysIle 562
XX :::::|||||
Db 2445 ---ACGTTTGGTGGACACACTGATGTGTGATCTTAAATAACAATTATACACATTTA 2501
QY 563 IleThrMetIleProAlaIleIleLysGlyAsnAsnLeuAspThrAsnSerLysValIleGlu 582
XX :::::|||||
Db 2502 ACTACCCAAATTCACGCTGTATAAAGGAATTCACCTTGGAGACGCTTCTAAGTTGTCCA 2561
QY 583 GlyProGlyHisThrGlyLysLeuValTyrLeuGlnSerGlnGlyArgLeuGluIle 602
XX :::::|||||
Db 2562 GGACCTGGTGCATACAGAGGAGGATTTAAT-----GATTTCAAAGATCATTTCAAAAT 2615
QY 603 ThrCysGluThrProAsnSerThrGlnSerTyrPheIleArgLeuAspGlyValThrAsn 622
XX :::::|||||
Db 2616 ACATGTCACACTCAAAATTTTCAACATCGATTTTATTAAGATTCGTTATCTTCAAAAT 2675
QY 623 GlyAlaGlyAsnThrLeuProAsnIleSerLeuThrIleProGlyValIleGlyIlePro 642
XX :::::|||||
Db 2676 GGAACGGCAAAATTAAGTCTGATCTTATTAATCTTAAATCCAGGGTACAGAACTG--- 2732
QY 643 ProGlnArgLeuAsnAsnThrPheSerGlyThrAsnTyrAsnAsnLeuGlnTyrGlyAsp 662
XX :::::|||||
Db 2733 GGTATGGCACTCAACCCCACTTTTCTGGTACAGATTATACGAATTTAAATATTAATAAT 2792
QY 663 PheGlyTyrPheGlnPheProSerThrValThrLeuProLeuAsnArgAsnIleProPhe 682
XX :::::|||||
Db 2793 TTTCACTTATTAAGATTTTCTAACAAGGTGAATTTGCTCCAAATCAAAACATCTCTT 2852
QY 683 IlePheAsnArgAlaAspVal---SerAsnSerIleLeuIleIleAspLysIleGluPhe 701
XX :::::|||||
Db 2853 GGTCTTAATCCGTTCCGATGTATATACAAACACACAGTACTTATGATTAATGAAATTT 2912
QY 702 IleProIleThrSerSerMetHisGlnAsnArgGlyLysGlnLysLeuGlnThrIleGln 721
XX :::::|||||
Db 2913 CTGCAATTAAGTCTGTTCTATAGAGAGATACAGAGAAACAATAATTAACAACAGTACAA 2972
QY 722 ThrLysIleAsnThrPhePheThrAsnHisThrLys 733
XX :::::|||||
Db 2973 CAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3008
XX :::::|||||

```

```

KW Insecticidal protein gene; Diptera; Aedes; ss.
XX
OS Bacillus thuringiensis serovar israelensis.
XX
FH Key Location/Qualifiers
FT CDS 1..4908
FT /tag= a
FT /product= "insecticidal protein"
FT Intron 620..1368
FT /tag= b
XX
XX JP63230090-A.
XX
XX 26-SEP-1988.
XX
XX 19-MAR-1987; 87JP-00066844.
XX
XX 19-MAR-1987; 87JP-00066844.
XX
XX (SDMO ) SUMITOMO CHEM IND KK.
XX
XX WPI; 1988-311968/44.
XX
XX P-PSDB; AAP81034, AAP82589.
XX
XX New insecticidal protein of bacillus thuringiensis israelensis strain -
XX prepd. by isolating plasmid contg. insecticidal protein gene by forming
XX gene library from plasmid deoxyribonucleic acid, etc.
XX
XX Disclosure; Page 7; 9pp; Japanese.
XX
XX A plasmid contg. the gene is isolated by forming a gene library from
XX CC plasmid DNA of B. thuringiensis israelensis (HD 522 strain, USA Goldberg
XX CC ONR60) followed by screening with anti-Israelensis insecticidal protein
XX CC IG. The insecticidal protein is highly effective against Diptera, esp.
XX CC Aedes. (Updated on 24-OCT-2003 to standardise OS field)
XX
SQ Sequence 4934 BP; 1797 A; 749 C; 877 G; 1511 T; 0 U; 0 Other;
XX
Alignment Scores:
Pred. No.: 4,196-85 Length: 4934
Score: 1059.50 Matches: 271
Percent Similarity: 50.44 Conservative: 118
Best Local Similarity: 35.18 Mismatches: 278
Query Match: 27.44 Indels: 105
DB: 1 Gaps: 24
XX
US-10-783-417-2 (1-735) x AAN81490 (1-4934)
QY 1 MetAsnGlnAsnAsnAspAsnAsnGluTyrGluIleIleAspSerHisThrSerProTyr 20
XX :::::|||||
Db 1369 ATGAATCCCTTATCAAAATTAATAATGAATATGAAACATTAAATGCTTCACAAAAAAATTA 1428
QY 21 PheProAsnArgAsnSerAsnAspSerArgTyrProTyrThrAsnAsnProAsnGlnPro 40
XX :::::|||||
Db 1429 -----AATATATCTCAATTAATTAATTAATCAAGATATCCAAATGAAATAGTCCAAACAAATTA 1482
QY 41 LeuGlnAsnThrAsnTyrLysGluThrPheLeuAsnMetCysGlnGlyAsnThrGlnTyrGly 60
XX :::::|||||
Db 1483 TTACAAAGTACCAATTAATTAAGATGCTCAATATGTGTCAACAGAAATACAGAGATGTGT 1542
QY 61 AsnAsnPheGluThrPheAlaSerAlaAspThrIleAlaIleValSerAlaGlyThrIle 80
XX :::::|||||
Db 1543 GGAAGATTTGAAACTTTTATGATAGT-----GGTGAACCTCAGTCCATATCTATT 1593
QY 81 ValSerGlyThrLeuLeuAlaGlyIleGlyLysLeuThrSerIleSerGlyProIleGly 100
XX :::::|||||
Db 1594 GTAGTTGGACCGTACTGACTGCTTCCGGTTCCACAACTCA-----CCCTTAGGA 1641
QY 101 IleIleGlyAlaIleIleSerPheGlyThrLeuIleThrValPheThrProAlaGly 120
XX :::::|||||
Db 1642 CTT-----GCTTAAATACGTTTGGTACATTAAACCGATCTTTTCCAGGCCAA 1692
QY 121 GluGlnAspLysThrValThrThrGlnPheIleLysMetGlyGluIlePheValAspThr 140
XX :::::|||||

```


JP2004166574-A.
17-JUN-2004.
19-NOV-2002; 2002JP-00335440.
19-NOV-2002; 2002JP-00335440.
(UYOK-) UNIV OKAYAMA.
WPI: 2004-445549/42.
Novel codon-modified gene encoding Cry4A protein, the modification being substitution with most frequently used codons in host microorganisms such that resulting gene has high translation rate in the microorganism.
Claim 3, SEQ ID NO 1; 20pp; Japanese.
The invention comprises a codon-modified gene which codes for Cry4A protein, the gene is modified by substitution with the most frequently used codon in the host microorganism, so that the modified gene will have a high translation rate in the host microorganism. The codon-modified gene of the invention is useful for the production of Cry4A protein in large quantities. The present DNA sequence represents a codon-modified Cry4A gene of the invention.
Sequence 2130 BP; 677 A; 704 C; 428 G; 321 T; 0 U; 0 Other;
Alignment Scores:
Pred. No.: 3,64e-83 Length: 2130
Score: 1032.50 Matches: 265
Percent Similarity: 50.3% Conservative: 117
Best Local Similarity: 34.9% Mismatches: 273
Query Match: 26.7% Indels: 105
DB: 12 Gaps: 24
US-10-783-417-2 (1-735) x ADP71294 (1-2130)
QY 1 MetAaNgInAaNsAaNsAaNsAaNgInUyRgIuIleIleAaSpSeRHiSeThSeRProTyR 20
DB 30 ATGAACCCGTCACCAAAACAGAAACGAAATACGAAACCTGACGACGACGACGACGACG 89
QY 21 PheProAaNsAaNsAaNsAaNsAaNsAaNsAaNsAaNsAaNsAaNsAaNsAaNsAaNsAa 40
DB 90 -----AACATCAGCAACAACTACACCCCTTACCCGATCGAAACAGCCGAAACACTG 143
QY 41 LeuGInAaNsAaNsAaNsAaNsAaNsAaNsAaNsAaNsAaNsAaNsAaNsAaNsAaNsAa 60
DB 144 CTGCAAGCAACCACTACCAAGACGCTGCAACATGTCACCAAAACCAACCAATACGCG 203
QY 61 AaPaenPheGluThrPheNlaSeRaIaAaPThrIleAlaIaValSeRaIaGlyThrIle 80
DB 204 GCGCACTTCGAACCTTATCGACAGC-----GGTGAACGTAGCCCGTACACATC 254
QY 81 ValSeRgIyThrLeuLeuAlaGlyIleGlyGlyLeuThrSerIleSerGlyProIleGly 100
DB 255 GTGGTCGGACCGTCTGACCGGTTCCGCTTCCACAC-----CCGCTGGGC 302
QY 101 IleIleGlyAlaIleIleIleSeRPhGlyThrLeuIleThrValPheThrProIaGly 120
DB 303 CTG-----GCGCTGATCGGTTCCGTTACCTGATCCGATCCGATCGGTCGCGCA 353
QY 121 GluGInAaPlyeThrValTrpThrGlnPheIleIleYsMeRgIyGluIlePheValaAaPThr 140
DB 354 GACCAAGCAACAC-----TGGAGCGACTTCATCACCACCAACCAAAACCTCATCAAAA 410
QY 141 ProLeuThrGlnSerIleIleGlnLeuIleGlnIleThrLeuGlnGlyPheArgGlnIle 160
DB 411 GAAATCGCAAGCACTTACATCAGCAACGCAACAAATCTCGAACCGTACGTTCAACGCG 470
QY 161 LeuGInSeRTrpAaNsThrAlaLeuAaPheTrpAaGlyIleValArgLeuGlnAlaPro 180
DB 471 ATCAAGCACTTACCAACCAACCTGAAAACCTGGAA-----AACAAACCG 515

QY 181 GlyLeuProProSeRSeRaIaLeuGlnIaIaIaLeuThrLeuIyIaArgPheGlu 200
DB 516 AACCGCAAAACACCAAGACGTGGTACCCAAATCCAACTGATGACCTTACCACTTCCAA 575
QY 201 AaNsValaIaAaNsAaPheIleArgGluIlePro-----GlyPheGlnLeuGlnThrTyR 218
DB 576 AACGTCACTCCCGCACTGTGAACAGCTGCCCGCAACCCGAGCGACTGCACTTAC 635
QY 219 LysThrLeuLeuProIleTyRAlaGlnIaIaIaAaPheIleIleAaNsLeuLeuGln 238
DB 636 AACATCTGTGTGAGCACTACCGCAAGCAAGCAACCTGCACTTACCGTCTGCAAC 695
QY 239 GlnGlyAlaGluLeuAlaAaPgluTrpAaNsIaAaPleIleIleProSeRgInIleGluPro 258
DB 696 CAAGCGCTCAAAATTCAGGCTTACCTGAAAACACCCCAATTCGACTTACCTGAGCCG 755
QY 259 AaNsAlaGlyThrSeRaAaPTrpTyRTrpIleLeuLeuIySgIuAaNsIleProIyTrSeR 278
DB 756 ---CTGCCGACCGCAATGACTTACCTCCGGTGTGACCAACGAAAGACTTAC 812
QY 279 AaNsTyRySaIaAaNsThrTyRArgThrGlyLeuIySgIuAaNsAaPgluPro----- 296
DB 813 AACCTGCTGAGCACTTACCAAAAGGCTTGAACCTGATCAAAACACCCGACAGC 872
QY 297 -----AaNsMetIySTrpSerIlePheAaNsAaPTrpArgTrpMetThrIle 312
DB 873 AACCTGACGCGCAACATCACTGAAACACTTACCAACACCTACCCGACCAAAATGACAGC 932
QY 313 ThrValLeuAaPThrIleSeRgInPheSeRLeuTyRAlaIleIySgIyTrpArgAaPSeR 332
DB 933 GCGGTGCTGAGCACTGTGTGCACTGTCCGCACTACGACGTCGTAATAC----- 983
QY 333 IleGlyGlyIleGluValIyGlyIleIySgIuAaNsIleuThrArgGluIleTyRThr 352
DB 984 -----CCGATCGGTGTCCAAAGCGACTGACCCGGAATCTTACCAATC 1028
QY 353 GluIleAaNsPheAaPArgLeuProGlnLeuArgValGlnProAaNsLeuAlaThrMetGlu 372
DB 1029 ---CTGAACCTTGAAGAAAGCCGTAACAAATATCTAC-----GACTTCAATATCAAGAA 1079
QY 373 TyrAaNsLeuThrArgAlaSeRPhLeuIySgIuAaNsPheIleGlnPheIleThrTyR 392
DB 1080 GACAGCCTGACCCGTCGCCG---CACTGTTCACCTGTGACAGCCTGAACCTTCTAC 1136
QY 393 -----ThrGluAaNsThraAaPheGlyAaNsAaG----- 401
DB 1137 GAAAAAGCGCAACCCCAACCACTTCTTACAGCAGCACTACAACTGTCACATC 1196
QY 402 ---LeuValaGlyIleSeRaAaNsAaPAlaProThrTyRSeAaNsThrIleThrGluThr 420
DB 1197 ACCCTGACCAACATCAGCAAAAGACGCGTTCGCAACCAACAGTGCAGCAAA 1256
QY 421 LeuTyRgIyGluAaGlyThrGlySeRProThrThrIySgIuAaNsProPheGluSeR 440
DB 1257 CTG-----AAAGCGTGTGTGCGCAACCAATCTTACATC----- 1292
QY 441 LysValSerIleValThraAaPgluSerProValSerProIleGlnProIlePhe 460
DB 1293 -----TTC 1295
QY 461 IleIleAaNsGlnIleGluLeu-----TyRLeuAaNsIySeSeRaAaNsThr--- 476
DB 1296 CTGCTGAACGTCATCAGCTGCAACAAATACCTGAAAGCACTTCAACACATCAGCAAA 1355
QY 477 -----LeuTyRTrSeRaIaGlyGly 483
DB 1356 ATGCACTTCTTATCAACCAACGTTACCGCTGTGAAAAAATCACTACGACAGCAGC 1415
QY 484 SerLeuSeRaAaNsTyRgIn---AaNsThrThrPhePheGlnPheProArgIyIySgIyAaP 502
DB 1416 GGCCAATCACTACGACGTGAACAAACATCTTGGCTGCGATCTGTAACGTCGC 1475

Oy	503	AenLeuVal11LeaSPProG1YCySeSerProaNPheAnaNPheAnTYSeSH1leLeuSer	522
Db	1476	GAaAACCAAGGCAACCCGACCTCTGTCGACCTACGCAACTACGCAACTCTTGAGC	1533
Oy	523	HisPheSerLeuPheThrTYrSerTYrVal11leGlyLeuGlnLeuGln11leuNaSPthr	542
Db	1536	TTTCATCAAAAGCGTGAAGCATCCGGGCAACCTACAAACCAAGTTC-----	1587
Oy	543	GlyValLeuGlyTYrTPHThiSeSerValaNPargTYrAaNa11leSerAPlysile	562
Db	1584	--ACCTTCGGCGTGAACCCACGACGCGCTCGACCCGAAAAAACCATCTACACCACTCG	1640
Oy	563	1leThrMet11leProAla11leGlyGlyAaNPheValaNPheAaNPheThraNPheSerTYrVal11leGln	582
Db	1641	ACCACCCAAATCCCGCGCGGTGAACGAACGACCTGGGCAACGCAAGCAAAAGTGTCCA	1700
Oy	583	GlyProGlyNH1STHGlYgLYAaNPheValTYrLeuGlnSerGlnGlyArgLeuGln1le	602
Db	1701	GGTCCGGGTCAACCGCGCGGTGACTGATC-----GACTTCAAAGACACTTCAAAATC	1755
Oy	603	ThrCysGlyIuThrProaNPheThrGlnSerTYrPhe11eArgLeuArgTYrAlaThraN	622
Db	1755	ACGTGCCAACACAGCAACTTCCAAACAAAGCTACTTCATCCGATCCGTTACGCGAGCAAC	1814
Oy	623	GlyAlaGlyAenThrLeuProaNPhe11leSerLeuThr11leProGlyVal11leGly1lePro	642
Db	1815	GGCAGCGGAAACACCCGCGCGGTGATCAACCTGACATCCCTCGGTGGCAGAACTG--	1871
Oy	643	ProGlnArgLeuAaNPheThrPheSerGlyThraNPheTYrAaNPheLeuGlnTYrGlyAaP	662
Db	1872	GGTATGGCACTGAACCCGACCTTCAGCGGTACCGACTACACCACTGAATCAAAAGAC	1931
Oy	663	PheGlyTYrPheGlnPheProSerThrValTYrLeuProLeuAaNPheAaNPheProPhe	682
Db	1932	TTCCAAATACCTGAAGTTCAAGCAACGAATGAATTCGGCGCGAACAATACTCAGCCTG	1991
Oy	683	11ePheAaNPheAlaNPheVal---SerAaNPhe11leLeu11leAaNPhe11leGluPhe	701
Db	1992	GTGTTCAACCGTAGGACGTGTACACCAACCAACCGTGCATCGACAAATCGAGTTTC	2051
Oy	702	11ePro11leThrSerSerMetNH1leGlnAaNPheGlyuysGlnTyLeuGlnThr11leGln	721
Db	2052	CTGCCGATCACCCGTAGCATCCGCGAAGACCGTGAAAAACAAACCTGAAACCGTGCA	2111
RESULT 10			
XX	ADP71295	ADP71295 standard; DNA; 2246 BP.	
AC	ADP71295;		
XX	09-SEP-2004	(first entry)	
DT			
XX			
DE		Codon-modified Cry4A gene #2.	
XX			
KM		codon optimisation; codon-modified; Cry4A; gene; ds.	
XX			
OS		Synthetic.	
XX		Unidentified.	
XX	JF2004166574-A.		
XX			
PD	17-JUN-2004.		
XX			
PF	19-NOV-2002; 2002JP-00335440.		
XX			
PR	19-NOV-2002; 2002JP-00335440.		
XX			
PA	(UYOK-) UNIV OKAYAMA.		
XX			
DR	WPI; 2004-445549/42.		
XX			
PT	Novel codon-modified gene encoding Cry4A protein, the modification being substitution with most frequently used codons in host microorganisms such		

Accession	Gene	Protein	Length	Score	Percent Similarity	Best Local Similarity	Query Match	DB
PF00001	that resulting gene has high translation rate in the microorganism.							
PS00001	Disclosure; SEQ ID NO 2; 20pp; Japanese.							
XX00001	The invention comprises a codon-modified gene which codes for Cry4A protein, the gene is modified by substitution with the most frequently used codon in the host microorganism, so that the modified gene will have a high translation rate in the host microorganism. The codon-modified gene of the invention is useful for the production of Cry4A protein in large quantities. The present DNA sequence represents a codon-modified Cry4A gene of the invention.							
XX00002	Sequence 2246 BP; 711 A; 728 C; 457 G; 350 T; 0 U; 0 Other;							
US-10-783-417-2 (1-735)	X ADP1295 (1-2246)							
QY	1 MetAsnGlnAsnAsnAsnAsnAsnGluTyrGluLeuLeuLeuAspSerHisThrSerProTyr	20	3,92e-83	2246				
DB	146 ATGAAACCCGTCACCAACGAAACGAAACGAAACCCGTCACCGAGCCGACGAAACGAAACGTCG	205	1032,50	265				
QY	21 PheProAsnAsnAsnSerAsnSerAsnSerProTyrProTyrThrAsnAsnProAsnGlnPro	40						
DB	206 -----AACATCGACCAACATCTACACCGCTTACCGATCGAAACAGCCGACGAAACGTCG	258						
QY	41 LeuGlnAsnThrAsnTyrLeuGluTyrPheLeuAsnMetCysGlnGlyAsnThrGlnTyrGly	60						
DB	260 CTCGAAAGCCCAACTACCAACAGCTGGCTGAACATGTGCCAACAACCAACATATACGCG	319						
QY	61 AspAsnPheGluThrPheAlaSerAlaAspThrIleAlaValSerAlaGlyThrIle	80						
DB	320 GGGCACTTCGAAACCTTCATCGACAGC-----GGTGAACAGCGCGCTACACCATC	370						
QY	81 ValSerGlyThrLeuLeuAlaGlyIleGlyLeuThrSerIleSerGlyProIleGly	100						
DB	371 GATGTCGACCGGTGTGACCGGTTTCGCTTACACAC-----CCGCTGAGC	418						
QY	101 IleIleGlyAlaIleIleIleSerPheGlyThrLeuIleThrValPheTrpProAlaGly	120						
DB	419 CTG-----CGCGTATCGGTTTCGCTACCTGATCCGCGGTCTGTTCCGCGCGCA	465						
QY	121 GluGlnAspLysThrValTrpThrGlnPheIleLysMetGlyGluIlePheValAspThr	140						
DB	470 GAGCAAGCAACG---TGAGCGACCTTCATCACCCCAACCAACAAATCATCATCAAAA	526						
QY	141 ProLeuThrGluSerIleLysGlnLeuLysLeuLeuThrLeuGluGlyPheAsnGlnIle	160						
DB	527 GAATATGCAAGACCTTACATCGAACCGCAACAAATCTCGAACCGTACGCTTCACGTC	586						
QY	161 LeuGlnSerTyrAsnThrAlaLeuAspAspTrpArgLysLeuLysValArgLeuGlnAlaPro	180						
DB	587 ATGACGACCTTACCAACACCTTGAACCTGGAA-----AACACCCG	631						
QY	181 GlyLeuProProSerSerAlaLeuGlnGlnAlaAlaLeuThrIleLysValArgPheGlu	200						
DB	632 AACCCGCAAAACACCAAGACGTCGTCACCAAAATCCAAATCGTGACCTACCACTTCMA	691						
QY	201 AsnValHisAsnAspPheIleArgGluIlePro-----GlyPheGlnLeuGluThrTyr	218						
DB	692 AACGTCATCCCGAATCGGTGAACAGCTGCCCGCGCAACCGGACGACCTGACCTACTAC	751						
QY	219 LysThrLeuLeuLeuProIleTyrAlaGlnAlaAlaAsnPheIleLysAsnLeuLeuGln	238						
DB	752 AACATCTCTGTCTGACGACGTACCGCGCAAGACGAACTCGACCTGACCGCTCTAAC	811						
QY	239 GlnGlyValGluLeuLeuAlaAspGluThrPheAlaAspIleHisProSerGlnIleGluPro	258						


```

Db      812  CAAAGGCTCAAAATTGAGGCTTACTGAAAAACAACCCGCAATTGACTACTCGACGCG 871
      Oy  259  AsnAlaGlyThrSerAspAspTyrTyrIleuLeuLysGluAsnIleProLysTyrSer 278
      Db  872  ---CGCCGACCGCATGACTACCTACCCGGTCTGACCAAAAGCAATGAGACTACACC 928
      Oy  279  AsnTyrCysAlaAsnThrTyrArgThrGlyLeuLysAsnLeuArgAspLysPro----- 296
      Db  929  AACTACTGCTGACCACTACAAAAAGCCCTGAACCTGATCAAAACACCCCGACAGC 988
      Oy  297  -----AsnMetLysTyrSerIlePheAsnAspTyrTyrArgTyrMetThrIle 312
      Db  989  AACCTGACGGCACAATCACTACGAAACACTACCAACCTACCGCACCAAAATGACCAAC 1048
      Oy  313  ThrValLeuAspThrIleSerGlnPheSerLeuTyrAspIleLysArgTyrArgAspSer 332
      Db  1049  GCGGTGCTGGACTGGCTGCTGTTCCGAACTACGACGTGCGTAAATAC----- 1099
      Oy  333  IleGlyGlyIleGluValLysGlyIleLysAsnGluLeuThrArgGluIleTyrThrThr 352
      Db  1100  -----CCGATCGGTGTCCAAAGCGAATCGAACCCGGGAATCTACCAAGTC 1144
      Oy  353  GluIleAsnPheAspArgLeuProGlnLeuArgValGlnProAsnLeuAlaThrMetGlu 372
      Db  1145  ---CTGAACCTTCGAAAGAAAGCCCGTACAAATACTAC-----GACTTCAATACCAAGAA 1195
      Oy  373  TyrAsnLeuThrArgAlaSerPheLysLeuPheSerPheLeuGlnGlnPheIlePheTyr 392
      Db  1196  GACAGCTGACCCGCTGCGCCG---CACTGTTCACTGCTGCTGACGCTGAACTTCTAC 1252
      Oy  393  -----ThrGluAsnThrAspAsnGlyAsnArg----- 401
      Db  1253  GAAAAAGCGCAACACCCCGAACAACCTTCTACACAGCACTACACAATGTTCCACTAC 1312
      Oy  402  ---LeuValGlyIleSerAsnArgAspAlaProThrTyrSerAsnThrIleThrGluThr 420
      Db  1313  ACCCTGGAACAACATGACCAAAAGACGCGTGTGCGACACCAACGTCGACCAACAA 1372
      Oy  421  LeuTyrGlyLysArgThrGlySerProThrThrLysThrIleArgProPheGluSerTyr 440
      Db  1373  CTG-----AAAGCCTGGGTCTGGCAACCAACATCTACATC----- 1408
      Oy  441  LysValSerIleValThrAspArgLysSerProValSerProIleGlnProHisPhe 460
      Db  1409  -----TTC 1411
      Oy  461  IleIleAsnGlnIleGluLeu-----TyrLeuAsnGlySerSerAsnAsnThr--- 476
      Db  1412  CTGCTGAACGTTCATCAGCTGACCAAAATACCTGAACGACTACACACATCTCAGCAA 1471
      Oy  477  -----LeuLysTyrSerLysGly 483
      Db  1472  ATGCACTTCTTCATACCAACGATACCGCGCTGCAAAAAAGAACTGACCGCAAGCAGC 1531
      Oy  484  SerLeuSerAsnTyrGln---AsnThrThrPhePheGlnPheProArgLysLysAspCys 502
      Db  1532  GGCCAATCACTGACAGTGTAACAAAAACATCTTCGGCTGCGCATCTGTAAGAGTGCGC 1591
      Oy  503  AsnLeuValIleAspProGlyCysSerProAsnPheAsnAsnTyrSerHisIleLeuSer 522
      Db  1592  GAAAAACAAGCAACCCGACCTGTTCCGACCTACGACCACTACAGCACTACCTCGAGC 1651
      Oy  523  HisPheSerLeuPheThrTyrSerTyrValIleGlyLeuGlnIleLeuAspThr 542
      Db  1652  TTCATCAAAAGCTGAGCTCCCGCAACTCAAAACCCCAAGTAC----- 1699
      Oy  543  GlyValLeuGlyTyrThrHisSerSerValAspArgTyrAsnAlaIleSerAspLysIle 562
      Db  1700  ---ACCTTGGGTGACCAAGCAGCGGTGCGCCGAAAAACACATCTACACCCACTG 1756
      Oy  563  IleThrMetIleProAlaIleLysGlyAsnAsnLeuAspThrAsnSerLysValIleGlu 582

```

```

Db      1757  ACCACCCAAATCCCGCGGTGAAAGCGAACAGCTGGCGACCGGACGAAAGGTGTCAA 1816
      Oy  583  GlyProGlyHisThrGlyGlyAsnLeuValTyrLeuGlnSerGlnGlyArgLeuGluIle 602
      Db  1817  GGTCCGGTTCACACCGCGGTGACCTGATC-----GACTTCAAAAGACACTTCAAAATC 1870
      Oy  603  ThrCysGluThrProAsnSerThrGlnSerTyrPheIleArgLeuArgTyrAlaThrAsn 622
      Db  1871  ACTGCAACACACAGCACTTCCAAAGACTTCACTCACTCCGATCCGTTACGCGAGCAAC 1930
      Oy  623  GlyAlaGlyAsnThrIleuProAsnIleSerLeuThrIleProGlyValIleGlyLysPro 642
      Db  1931  GCGACCGCAACACCCCGCGGTGATCACTGAGCATCCCTGGGTGCGAGAACTG--- 1987
      Oy  643  ProGlnArgLeuAsnAsnThrPheSerGlyThrAsnTyrAsnAsnLeuGlnTyrGlyAsp 662
      Db  1988  GGTATGGACTGAAACCCGACCTTACCGGTACCGACTACACCACTGAATACAAAGAC 2047
      Oy  663  PheGlyTyrPheGlnPheProSerThrValThrLeuProLeuAsnArgAsnIleProPhe 682
      Db  2048  TTCCAATACCTGAGATTGACGACAGAGTGAATTCGCGCGAACAACCAATCAGCCTG 2107
      Oy  683  IlePheAsnArgAlaAspVal---SerAsnSerIleLeuIleIleAspLysIleGluPhe 701
      Db  2108  GTGTTCAACCGGTAGCGAGGTGTACCAACACACCGGTGATCGACAAATCGAGTTC 2167
      Oy  702  IleProIleThrSerSerMetHisGlnAsnArgGlyLysGlnLysLeuGluThrIleGln 721
      Db  2168  CTGCGGATCACCCGTGACATCCGCGAAGACCTGTAAAAACAAAATCGGAACCTGTGCA 2227

RESULT 11
AA014670
ID   AA014670 standard; DNA; 2061 BP.
XX
XX
XX   AA014670;
XX
XX   27-AUG-2003 (revised)
XX   25-MAR-2003 (revised)
XX   04-FEB-1992 (first entry)
XX
XX   Dipteran active toxin gene.
XX
XX   Insecticide; B.c; crystal; delta endotoxin; cryIc; ss.
XX
XX   Bacillus thuringiensis serovar morileoni.
XX
XX   Key CDS
XX   1..2028
XX   /*tag= a
XX
XX   BP457498-A.
XX
XX   21-NOV-1991.
XX
XX   09-MAY-1991; 91EP-00304180.
XX
XX   15-MAY-1990; 90US-00524255.
XX   01-OCT-1990; 90US-00590903.
XX
XX   (MYCO ) MYCOGEN CORP.
XX
XX   Sick AJ;
XX
XX   WPI, 1991-341902/47.
XX   P-PSDB; AAR14374.
XX
XX   Bacillus thuringiensis genes encoding diptera-active toxins - and
XX   PT transformed microbes used to control insects in various environments.
XX   PS
XX   Clatm 1; Page 15; 20pp; English.
XX
XX   The sequence was obtd. from plasmid pMYC1636 which was isolated from a
XX   genomic library prepd. from DNA from B.c. PS71M3 [from B.c. PS71M3-69

```


QY 615 IleaTgLeuAryTYrAlaThraSnglYAlaGlYleuThrProAsnIleSerLeuThr 634
:::|||||
Db 1681 GTACGTATTCGTTAGTACTAATGCT-----CCAAAGCAACAGTATTC 1725
QY 635 IleProGlyValIleGlyIleProProGlnArgLeuAsnThrPheSerGlyThrAsn 654
:::|||||
Db 1726 TTAACCGGAATAGTACTATTAAGTGTGAG---CTCCCTAGTACCACTTCCGCCAAAC 1782
QY 655 Tyraen-----AsnLeuGlnTyrGlyAspPheGlyTyrPheGlnPheProSerThrVal 672
|||
Db 1783 CCAATGCTACAGATTTAACATATGCAGATTTTGGATATGTAACTTTCCAAAGAACGTT 1842
QY 673 -----ThreupProLeuAsnArgAsnIleProPheIle 683
|||
Db 1843 CCAATTAACAATTTGAAGAGAGAGACCTTATTAATGACC-----TTA 1887
QY 684 PheAsnArgAlaAspValSerAsnSerIleLeuIleIleAspValIleGluPheIlePro 703
:::|||||
Db 1888 TATGGTACACCAAAATCATTCAATATATATA--TATATTGACAAATGCAATTTATTC 1944
QY 704 IleThrSerSerMetHisGlnAsnArgGlyIleGlyIleGluThrIleGlnThrIle 723
|||
Db 1945 ATCACTCAATCTGTATTTAGATTATACAGAGAGCAAAATATAGAAACACAGAAATA 2004
QY 724 IleAsnThrPhePheThrAsnHisThrIleThrIleu 735
:::|||||
Db 2005 GTGAATGATTTATTTGTTAATTAACAAAGTTCTT 2040
RESULT 12
AA081180
ID AA081180 standard; DNA; 2061 BP.
AC AA081180;
XX
XX 25-MAR-2003 (revised)
DT 12-AUG-1995 (first entry)
XX
XX B.C. toxin PS71M3 gene.
DE
XX Delta-endotoxin; crystal protein; biological control agent; Calliphoridae;
KW screw-worm; sheep blowfly; Lucilia; Phormia; Calliphora; insecticide;
KM pesticide; B.C.; 86.
XX
XX Bacillus thuringiensis.
OS
XX MO9502694-A2.
PN
XX 26-JAN-1995.
PD
XX 13-JUL-1994; 94WO-US007902.
PE
XX 15-JUL-1993; 93US-00093199.
PR
XX (MYCO) MYCOGEN CORP.
PA
XX Hickie IA, Payne J;
PI
XX MPI; 1995-067338/09.
DR
XX P-P8DB; AAR63079.
PT
XX Method for controlling Calliphoridae pests - specifically utilises
PT Bacillus thuringiensis isolates or toxins.
XX
XX Disclosure; Page 42-43; 50pp; English.
XX
XX A library was constructed from Bacillus thuringiensis PS71M3 total
CC cellular DNA in lambda Gem-11. Plasmid pMYC1636, selected in Escherichia
CC coli, containing a 15 kb insert expressing a beta-endotoxin gene. This was
CC sequenced (AA081180). A cured, acrysaliferous B.t. host carrying
CC pMYC1636 produced a 77 kDa crystal protein (AAR63079). (Updated on 25-MAR
CC -2003 to correct PN field.)
XX

SEQ Sequence 2061 BP; 779 A; 331 C; 319 G; 632 T; 0 U; 0 Other;
Alignment Scores:
Pred. No.: 3,866-71 Length: 2061
Score: 899.50 Matches: 247
Percent Similarity: 48.38 Conservative: 126
Best Local Similarity: 32.08 Mismatches: 270
Query Match: 23.24 Indels: 129
DB: 2 Gaps: 30
US-10-783-417-2 (1-735) x AA081180 (1-2061)
QY 1 MetAsnGlnAsnAspAsnAsnGluTyrGluIleIleAspSerHisThrSerProTyr 20
|||
Db 1 ATGAATCCATATCAAAATAGAAATGAAATGAAATATTCATATGCTCCATCCAAATGTTT 60
QY 21 PheProAsnArgAsnSerAsnAspSerGlyTyrProTyrThrAsnAspProAsnGlnPro 40
|||
Db 61 -----ACCAAGTCAATTAATCTATTTGATATGATATTCATTAGCAAAATAGCCAAATCAACCA 114
QY 41 LeuGlnAsnThrAsnTyrGlyGluTyrPheAsnMetCysGlnGlyAsnThrGlnTyrGly 60
|||
Db 115 CTGAAAAACAGAAATTACAAAGATTGCTCAATGTGTCTCAAGATATCAACAAATATGCT 174
QY 61 AspAsnPheGluThrPheAlaSerAlaAspThrIleAlaAlaValSerAlaGlyThrIle 80
|||
Db 175 AATAATGCGGGGAATTTTGTAGTTCTGAACATATGTTGGAGTTAGTACAGGTATATT 234
QY 81 ValSerGlyThrLeuLeuAlaGlyIleGlyGlyLeuThrSerIleSerGlyProIleGly 100
|||
Db 235 GTAGTAGGAATCACTATGATGGA-----GCTTTGCTGCCCT----- 270
QY 101 IleIleGlyAlaIleIleIleSerPheGlyThrIleuIleThrValPheTyrProAlaGly 120
|||
Db 271 GTCTTAGCTGACGATATATATCTTTGGACCTTTGTCGCCGATCTTTGG---CAAGCA 327
QY 121 GluGlnAspLysTrpValTyrThrGlnPheIleLysMetGlyGluIlePheValAspThr 140
|||
Db 328 TCTGACCCCTGCAAATGTTTGGCAGGATTTGTAACATGCA-----GGAAG 375
QY 141 ProLeuThrGlu-----SerIleGlyGlnLeuLysLeuGlnThrLeuGlnGlyPhe 157
|||
Db 376 CTTATACAGAAATAGATTAACAACTAATTAATGACTTAACCTTCTATGTAACACTATA 435
QY 158 ArgGlnIleLeuGlnSerTyrAsnThrAlaLeuAspAspTyrPargLysLeuLysArgLeu 177
|||
Db 436 AAAAATCACTTGATTAATATCAAGAATTTTGCATTAATGGACCCAGCAAGT----- 489
QY 178 GlnAlaProGlyLeuProProSerSerAlaLeuGlnGlnAlaAlaLeuThrLeuLysIle 197
|||
Db 490 -----ACACACGCTAATGCTAAAGCAGTACATGATCTTTACT 528
QY 198 ArgPheGluAsnValHisAsn-----AspPheIleArgGluIleProGlyPheGln 214
|||
Db 529 ACCTTAGAACCTATTAATAGATTAAGATTATGTTAAATAATATGCTAGCATATGCA 588
QY 215 LeuGluThrTyrLysThrLeuLeuProIleTyrAlaGlnAlaAlaAsnPheHisLeu 234
|||
Db 589 ATACCAACA-----CTCCCTCATATGCAACAATAGCTACTTGGACACTTG 633
QY 235 AsnLeuLeuGlnGlnGlyAlaGluLeuAlaAspGluTyrPheAlaAspIleHisProSer 254
|||
Db 634 AATTTATTAACATGCTGCTACTATTAACATATATG-----CTGCAAAATCA 684
QY 255 GlnIleGluProAsnAlaGlyThrSerAspAspTyrTyrLys---LeuLeuGluAsn 273
|||
Db 685 GGTATTAATCCAAAGTACTTCAATTCATCTAATTAATCAAGGGCTATTTAACAGTAAA 744
QY 274 IleProLysTyrSerAsnTyrCysAlaAsnThrTyrArgThrGlyLeuLysAsnLeuArg 293
|||
Db 745 ATACAAAGATATCTGACTATTTGTATACAAAGTACAAAGTACGAGACTAATATGATTTGA 804
QY 294 AspGluProAsnMetLysTyrSerIlePheAsnAspTyrArgArgTyrMetThrIleThr 313


```
US-10-783-417-2 (1-735) x AAN50525 (1-3756)
OY 1 MetAnglnAsnAspAsnAsnGluYrGluIleIleAspSerHisThrSerProTyr 20
DB 941 ATGAATCCATATCAAAAATAGAAATGAAATGAAATTTCAATGCTCCATCCAAATGGTTT 1000
OY 21 PheProAsnArgAsnSerAsnAspSerArgTyrProTyrThrAsnAsnProAsnGlnPro 40
DB 1001 -----AGCAAGCTAAATATCTATTCTAGATATCCATTAGCAAAATAGCCAAATCAACA 1054
OY 41 LeuGlnAsnThrAsnTyrIleGluTyrPheAsnMetCysGlnGlyAsnThrGlnTyrGly 60
DB 1055 CTGAAAACACAGAAATTCACAAAGATTGGCTCAATGCTGTGCAAGATATCAACAATATGCG 1114
OY 61 AspAsnPheGluThrPheAlaSerAlaAspThrIleAlaIleValSerAlaGlyThrIle 80
DB 1115 AATTAATGGCGGGAATTTCTGACTGCTGAACTATTGTTGGAGTTAGTCAAGTATATT 1174
OY 81 ValSerGlyThrLeuLeuAlaGlyIleGlyGlyLeuThrSerIleSerGlyProIleGly 100
DB 1175 GTAGTAGGAATCTATGTTAGCA-----GCTTTTGCTGCCCT----- 1210
OY 101 IleIleGlyAlaIleIleIleSerPheGlyThrLeuIleThrValPheTyrProAlaGly 120
DB 1211 GCTTAGCTGACAGATATATATCTTTGGACCTTGTTGCCGATCTTTGG--CAAGCA 1267
OY 121 GluGlnAspLeuThrValTyrThrGlnPheIleLeuMetGlyGluIlePheValAspThr 140
DB 1268 TCTGACCCCTGCAAAATGTTGGCAGATTTGTTAAACATCGA-----GGAAG 1315
OY 141 ProLeuThrGlu-----SerIleGlyGlnLeuLeuYleuGlnThrLeuGlnGlyPhe 157
DB 1316 CCTATGCAAGAAATAGATATAAAACATTAATGATGACTTAATCTTCTGTAACACTTANA 1375
OY 158 ArgGlnIleLeuGlnSerTyrAsnThrAlaLeuAspAspTyrArgIleLeuLeuArgLeu 177
DB 1376 AAAAATCAACTGATTAATATCAAGAAATTTTCGATAAATGGAGCCACACAGT----- 1429
OY 178 GlnAlaProGlyLeuProSerSerAlaLeuGlnGlnAlaAlaLeuThrLeuYsIle 197
DB 1430 -----ACACACGCTAATGCTAAAGCAGTACATATCTTTTACT 1468
OY 198 ArgPheGluAsnValHisAsn-----AspPheIleArgGluIleProGlyPheGln 214
DB 1469 ACCTTAGAAGCTATATAGATTAAGATTAGATATGTTAAAAAATAGCTAGCTATGCA 1528
OY 215 LeuGluThrTyrIleThrLeuLeuLeuProIleTyrAlaGlnAlaAlaAsnPheHisPhe 234
DB 1529 ATACCAACA-----CTCCCTGCATATGCAACAATAGCTACTTGGCAGCTTG 1573
OY 235 AsnLeuLeuGlnGlnGlyValaGluLeuAlaAspGluTyrPheAsnAlaAspIleHisProSer 254
DB 1574 AATTATATTAACATGCTGCTACTCTTCAATTTCACAAATATATG-----CTGCAAAATCA 1624
OY 255 GlnIleGluProAsnAlaGlyThrSerAspAspTyrTyrIle-----LeuLeuYsGluAsn 273
DB 1625 GGATTAATATCCAAAGTCTTCAATTCATCTAATTAATCTATCAGGGCATTTAAAGGTAA 1684
OY 274 IleProIleTyrSerAsnTyrCysAlaAsnThrTyrArgThrGlyLeuLeuAsnLeuArg 293
DB 1685 ATACAGAAATATATGCTACTATGTATACAAACGTCAATGCAAGTACTATGATTTTAA 1744
OY 294 AspGluProAsnMetLeuTyrSerIlePheAsnAspTyrArgArgTyrMetThrIleThr 313
DB 1745 ACTAATATCTAACGCAACATGAATATGATTAATACTTACCGTTTGAATAAGCTCTAAT 1804
OY 314 ValLeuAspThrIleSerGlnPheSerLeuTyrAspIleLeuArgTyrArgAspSerIle 333
DB 1805 GGTATGATCTTATGCTATTTTTCGAATTAATGACCCGAGAAAATAT----- 1852
OY 334 GlyGlyIleGluValIleGlyIleYsAsnGluLeuThrArgGluIleTyrThrThrGlu 353
DB 1853 -----CCAAATGAGATTAAATCTGAACTTATCAGAGAACTTAT--ACGAAT 1897
OY 354 IleAsnPheAspArgLeuProGlnLeuArgValGlnProAsnLeuAlaThrMetGluTyr 373
DB 1898 GTTAATTCAGATACATTT-----AGAACATATAACGAATCTAGAAAAT 1939
OY 374 AsnLeuThrArgAlaSerPheLeuLeuPheSerPheLeuGlnGlnPheIlePheTyrThr 393
DB 1940 GGATTAATCTAG--AATCTTACATTAATTAATCTTGATTAACCAAGGCGCTTTTACACA 1996
OY 394 GluAsnThrAsn-----PheGlyAsnArgLeu 402
DB 1997 AGAAATTCCTCAGACATCTTGATCTTAATGATATTTTCTTTTACAGTAACAGATG 2056
OY 403 ValGlyIleSerAsnArgAspAlaProThrTyrSerAsnThrIleThrGluThrLeuTyr 422
DB 2057 -----GCTTTACACATTAATTAATGATGCGCAACATATCTGG 2095
OY 423 GlyIleArgThrGly-----SerProThrThrIleThrIleArgProPheGluSer 439
DB 2096 GGAGCGGTTCAATGAAATATTAATTTCTCAGACACATCCAAAGTATTTCTTTTATAGA 2155
OY 440 Tyr-----IleValSerIleValThrAspArgGlnSerProProValSerPro 455
DB 2156 AACAACTTATGATTAAGTCAAAATTCATGACATTAAGAGTACTCAGATTAATA-- 2212
OY 456 IleGlnProHisPheIleIleAsnGlnIleGluLeuTyrLeuAsnGlySerSerAsn 475
DB 2213 -----TATGAAATGATATTTTTCGAATAGCAGTGA 2245
OY 476 ThrLeuYsTyrSerAlaGlyIleSerLeu-----SerAsnTyrGlnAsnThrThrPhePhe 494
DB 2246 GTATTTCATATTCATCAACATTCACAAATAGAAATAATTAATAAAGAACTGATCTTAT 2305
OY 495 GlnPheProArgGlyLeuAspCysAsnLeuValIleAspProGlyCysSerProAsnPhe 514
DB 2306 ATGATTCCAAACCAACATGGAATAAT----- 2332
OY 515 AsnAsnTyrSerHisIleLeuSerHisPheSerLeuPheThrTyrSerTyrValIleGly 534
DB 2333 CAAGAAATATGCTACATCTATCTATATTAATAAATCTGATTAATTAATTTCACTAGT 2392
OY 535 LeuGlnLeuGlnIleLeuAspThrGlyValLeuGlyTyrPheHisSerSerValAspArg 554
DB 2393 AGAGAAAGAAAGAGAGT-----GCATTAGTGGACACATATGATGTTGATTTTC 2443
OY 555 TyrAsnAlaIleSerAspLeuIleIleThrMetIleProAlaIleIleGlyYsAsnLeu 574
DB 2444 CAAAATACAAATGATTTAGATTAACATCAACCAATCCACGCTCTAATAAATCTTAAAGTA 2503
OY 575 AspThrAsnSerLeuValIleGluGlyProGlyHisThrGlyGlyYsAsnLeuValTyrLeu 594
DB 2504 AGTTCTAATTCAAAATTTGTGAAGCGCTCGGTCAACAAGAGAAATCTGTAATCTT 2563
OY 595 GlnSerGlnGlyArgLeuGluIleThrCysGluThrProAsnSerThrGlnSerTyrPhe 614
DB 2564 AAAAGATAGATGAATTTAGAGTTAGATT--TAAAAAAGTTTCCACAATAATCA 2620
OY 615 IleArgLeuArgTyrAlaIleThrAsnGlyAlaGlyAsnThr-----LeuPro 629
DB 2621 GTACGTAATTCGTAATGCTAATGCTCCAAAGACACAGATTTCTTAAACCGGAATAGAT 2680
OY 630 AsnIleSerLeuThrIleProGlyValIleGlyIleProProGlnArgLeuAsnAsnThr 649
DB 2681 ACTATTAAGTCTGAGCTCCT-----AGTACC 2707
OY 650 PheSerGlyThrAsnTyrAsn-----AsnLeuGlnTyrGlyAspPheGlyTyrPheGln 667
DB 2708 ACTCCCGCAAAACCAAAATGCTACAGATTTTAATATATGACGATTTTGGATATAGAAC 2767
OY 668 PheProSerThrValThrLeuProLeuAsnArgAsnIle-----ProPheIle 683
DB 2768 TTTTCAGACAGCTTCA-----AATTAACATTTGAAGAGAGACACTTTAT 2818
```

QY 684 -----PheaaNrgAlaapValSerAaSerileleuileleaplysile 699
DB 2819 AATGACCTTATTAATGTAACCAAAATCTTCAATATAATTA---TATTTGACAAAT 2875
QY 700 GlupheileProileTherSerMerethiagInaasnArgglulysGlnlyleugluthr 719
DB 2876 GAATTTATTCATCACTCAATCTGTATTAAGATTATACAGAGAACAAATATAGAAA 2935
QY 720 IleglnthrlvleasnthrPhePethrAaenHlsthlyleThleu 735
DB 2936 ACACAGAAATAGTGAATTTATTTATTAATAAACAAAGTTCTT 2983
RESULT 14
AAD43974
ID AAD43974 standard, DNA; 4896 BP.
XX
AC AAD43974;
XX
DT 13-DEC-2002 (first entry)
XX
DE Bacillus thuringiensis ssp. finitimus cry28Aa1 gene.
XX
KM Delta-endotoxin; cry26Aa1; insect-resistant plant; toxin;
XX transgenic host cell; insecticide; gene; ds.
OS Bacillus thuringiensis.
XX
XX Key Location/Qualifiers
FH CDS 1129..4458
FT /tag= a
FT /product= "cry28Aa1 gene"
XX
XX US2002038005-A1.
XX
XX 28-MAR-2002.
XX
XX 08-JAN-2001; 2001US-00756526.
XX
XX 07-JAN-2000; 2000US-0175158P.
XX
XX (WOJC/) WOJCIECHOWSKA J A.
XX (LEWI/) LEWITTIN E I.
XX (ZALU/) ZALUMIN I A.
XX (REVI/) REVINA L P.
XX (CHES/) CHESTUKHINA G G.
XX
XX Wojciechowska JA, Lewitin EI, Zalumin IA, Revina LP;
XX Chestukhina GG;
XX
XX WPI; 2002-403936/43.
XX P-PSDB; AAE26353.
XX
XX Novel isolated delta-endotoxin nucleic acid molecules, cry26Aa1 and
XX cry28Aa1 isolated from Bacillus thuringiensis finitimus, that encodes
XX toxin active against insects, useful for controlling insects.
XX
XX Claim 1; Page 33-37; 42pp; English.
XX
XX The invention relates to isolated delta-endotoxin nucleic acid molecules,
XX cry26Aa1 and cry28Aa1 isolated from Bacillus thuringiensis finitimus,
XX that encode a toxin that is active against insects. The invention is
XX useful for producing an insect-resistant plant, by introducing the
XX nucleic acid molecule into the plant, where the nucleic acid is
XX expressible in the plant in an effective amount to control an insect. The
XX invention is useful for producing a toxin that is active against insects
XX by obtaining the transgenic host cell and expressing the nucleic acid
XX molecule in the host cell, which results in the toxin that is active
XX against insects. The toxin is useful for controlling an insect by
XX delivering to the insect an effective amount of toxin. The invention is
XX useful for controlling insects. The toxin is useful for inhibiting the
XX ability of insect pest to survive, grow or reproduce, for limiting insect
XX-related damage or loss in crop plants, and to prophylactically treat

CC Insect susceptible areas or plants to confer protection or resistance
CC against harmful insects. The present sequence is Bacillus thuringiensis
CC ssp. finitimus cry28Aa1 gene
XX
SQ Sequence 4896 BP; 1727 A; 757 C; 887 G; 1525 T; 0 U; 0 Other;
Alignment Scores:
Pred. No.: 4.8e-66 Length: 4896
Score: 849.00 Matches: 229
Percent Similarity: 48.8% Conservative: 106
Best Local Similarity: 33.3% Mismatches: 248
Query Match: 21.9% Indels: 104
DB: Gaps: 25
US-10-783-417-2 (1-735) x AAD43974 (1-4896)
QY 68 SerAlaapThrIleAlaIaValSerAlaGlyThrIleValSerGlyThrLeuLeuAla 87
DB 1210 AGCAGTGATACAGTCGTGTAGTAAGCCAGGAGATTGTGGGTACTATACAGACA 1269
QY 88 GlyIleGlyGlyLeuThrSerIleSerGlyProIleGlyIleIleGlyAlaIleIleIle 107
DB 1270 -----GCCTTGCATTCATTGTATATCA-----GGTGTGACTATATA 1308
QY 108 SerPheGlyThrLeuIleThrValPheTrpProAlaGlyGlnGlnAsp---LysThrVal 126
DB 1309 TCATTTGAACTTGTGCTCCGCTTCCTTGCTGATCCAGAGAAATCCAAAAAATT 1368
QY 127 TrpThrGlnPheIleLysMetGlyGlnIlePheValAspThrProLeuThrGlnSerIle 146
DB 1369 TGGTCACATTTATGAAACACGAGAGACCTTTAAATCAAAATTTCTACAGCTGTA 1428
QY 147 LysGlnLeuLysLeuGlnThrLeuGlnLysPheArgGlnIleLeuGlnSerIleThr 166
DB 1429 AAAAGAAATGCAATTCATTCATTAATGTTTAAAGTATTAACGTACTATGAAGA 1488
QY 167 AlaLeuAspAspTrpArgGlyLysLeuLysArgLeuGlnAlaProGlyLeuProSerSer 186
DB 1489 GCATTTAATGATTTGAAGAAGA-----AATCCAACTGCA 1521
QY 187 AlaLeuGlnGlnAlaIleLeuThrLeuLysIleArgPheGlnAsnValHisAsnAspHe 206
DB 1522 -----AATACGCCAGATGGTATGATACAGAGATTTGAAAGCCTCATTTCAATTT 1572
QY 207 IleArgGlnIleProGlyPheGlnLeuGlnThrIleLysThrLeuLeuLeuProIleIle 226
DB 1573 GTACCAATATGCCCAACTCCCACTCCACGTATGACACATTTATTAATGTTGCTAT 1632
QY 227 AlaGlnAlaIaAsnPheHisLeuAsnLeuLeuGlnGlnIleValGlnLeuAlaAspGlu 246
DB 1633 ACAGAACTGCAAAATTTACATTTGAATTTATTAATCAAGGTGTAACAATTCGCGGATCA 1692
QY 247 TrpAsnAlaAspIleHisProSerGlnIleGlnProAsnAlaGlyThrSerAspAspIle 266
DB 1693 TCGAATGCAAGTCAACACATTTCAACAAATGTTGAAGTATCATCGATCACT-----TAT 1743
QY 267 TyrLysLeuLeuLysGlnAsnIleProLysIleYserAsnTrpCysAlaAsnThrIleArg 286
DB 1744 TATGACGAGCTATGTGTATATTTGAAAAAGTATTAATTTATGCAACCAAGCATTCAT 1803
QY 287 ThrGlyLeuLysAsnLeuArgAspGlnProAsnMetLysTrpSerIlePheAsnAspIle 306
DB 1804 AAAGATTTGAATCACCTTAAAGAAATCAAGAAAAATCAACATGGAGATCTTATACACAT 1863
QY 307 ArgArgTrpMetThrIleThrValLeuAspThrIleSerGlnPheSerLeuIleAspIle 326
DB 1864 CGTCGAGAAATATACCTTATATGTATGATCTGTGCAACTTTCCTTTATATATATA 1923
QY 327 LysArgTrpArgAspSerIleGlyGlyIleGlnValLysGlyIleLysAsnGlnLeuThr 346
DB 1924 CCTCCTTTT-----CCAGAGGAGTGAATCAATTAATTAACA 1959
QY 347 ArgGlnIleIleThrThrGlnIleAsnPheAspArgLeuProGlnLeuArgValGlnPro 366


```

Db      1960 AGAGAGGTTTATACAGTTTATAGATCATTTAAACACACACCA----- 2001
Qy      367 AenleuAlaThrMetGluTyrAsnleuThrArgAlaSerPheLeuPheSerPheLeu 386
Db      2002 -----GGGCTATTATCTGGCTG 2019
Qy      387 GluGlnPheIlePheTyrThrGluAsnThrAsnPheGlyAsnArgLeuValGlyIleSer 406
Db      2020 TCAGATATTGAGTTTACACGAGAGGTGGCAGAAAGCAGATTATTTATCAGTATTT--- 2076
Qy      407 AsnArgAlaProThrTyrSerAsnThr-----IleThrGluThrLeuTyrGly 423
Db      2077 ---CGAGAGTCTAAATATTATTAAGTGTATCAATTTTATACAGTAAATATTATGCT 2133
Qy      424 GluArgThrGly---SerProThrThrIleArgProPheGluSerTyrLeuVal 442
Db      2134 AATACAAATAGATTAGTAAAGAGCTCATTTATTTACCGAGGAA---TTTATGACT 2190
Qy      443 SerIleValThrAspArgGlnSerProProValSerProIleGlnProIlePheIleIle 462
Db      2191 CACTTAAGCATVAAACGTCCTTTCAAAACAATAGCTGGATTAATTAAGTTATACAGTTTA 2250
Qy      463 AsnGlnIleGluLeuTyrLeuAsnGlySerSerAsnAsnThrLeuTyrSerAlaGly 482
Db      2251 ATTCAAAAAATCGTATTCACAACTTTTAAAAACGATAT---GAATATCAAAAAAAT 2304
Qy      483 GlySerLeuSerAsnTyr-----GlnAsnThrThrPhePheGlnPheProArgLys 499
Db      2305 TTTATGTGATTAATCAAAATGACACCTCAAGAACTTACA----- 2343
Qy      500 LysAspCysAsnLeuValIleAspProGlyCysSerProAsnPheAsnAsnTyrSerHis 519
Db      2344 -----AAGTATCTTAATGATATGATGTTGTTCA---AACAGCCAAAAATTAACAAT 2391
Qy      520 IleLeuSerHisPheSerLeuPheThrTyrSerTyrValIleGlyLeuGlnLeuGlnIle 539
Db      2392 AATTATCTCATTTTCCATTTATCATCCAC-----AAGTTACAGTTT 2433
Qy      540 LeuAsp-----ThrGlyValLeuGlyTyrThrHisSerSerValAspArgTyr 555
Db      2434 GCTGAGTATTTTCACTCATATTTTGCATTAGGTTGAGCACACAAATAGTGAATCCCA 2493
Qy      556 AsnAlaIleSerAspLysIleIleThrMetIleProAlaIleIleGlyAsnAsnLeuAsp 575
Db      2494 AATTAAATATCAGAAAGGTGAGTACACAAATCCATTGGTAAAGCTTACGAAGTT--- 2550
Qy      576 ThrAsnSerLysValIleGlyIleProGlyHisThrGlyGlyAsnLeuValTyrLeuGln 595
Db      2551 ACTAACAAATTCAGTTATTAAGAGACAGGTTTACAGGTGAGATTTAATGAACCTTCGT 2610
Qy      596 SerGlnGlyArgLeuGluIleThrCysGluThrProAsnSerThrGlnSerTyrPheIle 615
Db      2611 GAT-----AAATGTTCTATTAAAGTAAAGCT---AGTCTTTAAAAAAAATCCCTATA 2661
Qy      616 ArgLeuArgTyrAlaThrAsnGlyValGlyAsnThrLeuProAsnIleSerLeuThrIle 635
Db      2662 AGTCATTTTATGCTGCAAAATACGCAATAGCTGATCAATAGACGTAGATTCGCGA 2721
Qy      636 ProGlyValIleGlyIleProProGlnArgLeuAsnAsnThrPheSerGlyThrAsnTyr 655
Db      2722 GCAGAGAGTCTA-----TTGCAACTACCTTTTCTAGAAAAGGGAAC 2763
Qy      656 AsnAsn-----LeuGlnTyrGlyAspPheGlyTyrPheGlnPheProSer 670
Db      2764 AATAAATTTTACAAATCAAGACCTTAACATTAAGATTTTCAATATCATACACTTTAGCT 2823
Qy      671 ThrValThrLeuProLeuAsnArgAsnIleProPheIlePheAsnArgAlaAsp----- 688
Db      2824 GATATTGAATTATCCGAAAGATGAAGAATTCATATCCATTGAACGAGAGATGATAT 2883
Qy      689 ValSerAsnSerIleLeuIleIleAspLysIleGluPheIleProIleThrSerSerMet 708

```

```

Db      2884 GAGAGGAGTGAATCTTTTATGATTAATTGATTAAGATTCAAACTTATGATGAAAAATTAT 2943
Qy      709 HisGlnAsnArgGlyLeuGlnLeuGluThrIleGlnThrIleAsnThrPhePhe 728
Db      2944 -----ACTAATGAATAATTAATTAAGAGGCAAGCAAGCAAGTGAATGTTATTT 2994
Qy      729 ThrAsnHisThrLysThrLeu 735
Db      2995 ATAAACGCAACAAACGCTTTG 3015

RESULT 15
ADFI3101
ID ADFI3101 standard; DNA; 4896 BP.
AC ADFI3101;
DT 12-FEB-2004 (first entry)
XX
DE Bacillus thuringiensis serovar finitimus pf2 DNA clone.
XX
KW Cry26Aa1; cry28Aa1; delta-endotoxin; insect pest control;
XX transgenic plant; insect resistance; insecticide; gene; ds.
XX
OS Bacillus thuringiensis serovar finitimus.
XX
FH Key Location/Qualifiers
FH CDS 1129..4458
FT /*tag= a
FT /product= "Cry28Aa1 delta-endotoxin protein"

XX
EN US2003150018-A1.
XX
PD 07-AUG-2003.
XX
PE 15-JAN-2003; 2003US-00345020.
XX
PR 07-JAN-2000; 2000US-0175158P.
PR 08-JAN-2001; 2001US-00756526.
XX
PA (WOJCI/) WOJCIECHOWSKA J A.
PA (LEMIT/) LEWITIN E I.
PA (ZALUN/) ZALUNIN I A.
PA (REV1/) REVINA L P.
PA (CHES/) CHESTUKHINA G G.
XX
PI Wojciechowska JA, Lewitin EI, Zalunin IA, Revina LP;
PI Chestukhina GG;
XX
DR WP1; 2003-897623/82.
DR P-PSDB; ADFI3102.
XX
PT New isolated nucleic acid molecule encoding a toxin that is active
PT against insects useful for controlling insect pests or for conferring
PT insect resistance in plants.
XX
PS Claim 7; SEQ ID NO 3; 42pp; English.
XX
CC The present invention relates to the isolation of novel cry26Aa1 and
CC cry28Aa1 delta-endotoxin genes from Bacillus thuringiensis sep.
CC finitimus. The sequences for the delta-endotoxin polypeptides are also
CC disclosed. The invention provides methods for producing the toxins and
CC compositions containing the toxins. The methods and sequences of the
CC invention are useful for controlling insect pests in transgenic plants to
CC confer insect resistance. The present sequence represents a DNA clone
CC that contains the coding sequence for cry28Aa1 delta-endotoxin.
XX
SQ Sequence 4896 BP; 1727 A; 757 C; 887 G; 1525 T; 0 U; 0 Other;
XX

Alignment Scores:
Pred. No.: 4.8e-66 Length: 4896
Score: 849.00 Matches: 229
Percent Similarity: 48.8% Conservative: 106
Best Local Similarity: 33.3% Mismatches: 248

```

Query Match: 21.9% Indels: 104
DB: 10 Gaps: 25
US-10-783-417-2 (1-735) x ADFJ1301 (1-4896)
QY 68 SerAlaAspThrIleAlaAlaValSerAlaGlyThrIleValSerGlyThrLeuLeuAla 87
DB 1210 ACCAGATATACAGTCTGCTAGTAAAGCAGGAGATGTTGGTCTACTATACAGACA 1269
QY 88 GlyIleGlyGlyLeuThrSerIleSerGlyProIleGlyIleIleGlyAlaIleIleIle 107
DB 1270 -----GCTTTCATCATTTTGTAAATCCA-----GCTGTGACTTAA 1308
QY 108 SerPheGlyThrLeuIleThrValPheTTPProAlaGlyGluGlnAsp---LysThrVal 126
DB 1309 TCATTGGAACTTGGCTCCGTTCTTGGCTGATCCAGAGAAATCCAAAATAAT 1368
QY 127 TTPThrGlnPheIleLysMetGlyGluIlePheValAspThrProLeuThrGluSerIle 146
DB 1369 TGGTCAATTTATGAAACACGAGAAACCTTTAAATCAAAATTTCTACAGCTGA 1428
QY 147 LysGlnLeuLysLeuGlnThrLeuGluGlyPheAspGlnIleLeuGlnSerTyraThr 166
DB 1429 AAGAATATGATTCATACCTCAATTAATGTTTAAAGATGATTAAGTACTATGAAGA 1488
QY 167 AlaLeuAspAspThrArgLysLeuLysArgLeuGlnAlaProGlyLeuProProSerSer 186
DB 1489 GCATTTAATGATTTGAAGAGA-----AATCCAGCTGA 1521
QY 187 AlaLeuGlnAlaAlaLeuThrLeuLysIleArgPheGluAsnValHisAsnAspPhe 206
DB 1522 -----AATATCCAGATTTGTATGACAGAGATTTTGAAGACCTCATTTCAATTT 1572
QY 207 IleArgGluIleProGlyPheGlnLeuGluThrTyLysThrLeuLeuLeuProIleTy 226
DB 1573 GTAGAGATATGCAACACTCCAACTCCACGATGACACATTAATTAAGTTCTAT 1632
QY 227 AlaGlnAlaAlaAsnPheHisLeuAsnLeuLeuGlnGlyAlaGluLeuAlaAspGlu 246
DB 1633 ACAGAGCTGCAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1692
QY 247 TrpAsnAlaAspIleHisProSerGlnIleGluProAsnAlaGlyThrSerAspAspTy 266
DB 1693 TGGATATGATATCAACATTCACCAATGTTGAATCATCAAGTACT-----TAT 1743
QY 267 TyLysLeuLeuLysGluAsnIleProLysTySerAsnTyCysAlaAsnThrTyArg 286
DB 1744 TATGACGACTATTTGGTATATATGAAAGTATATTAATTAATTAATTAATTAATTA 1803
QY 287 ThrGlyLeuLysAsnLeuArgAspGluProAsnMetLysTrpSerIlePheAsnAspTy 306
DB 1804 AAGAATTAATCACTTAAGAAATCAAGAAATAATCAATGGATGCTTAATTAACATAT 1863
QY 307 ArgArgTyMetThrIleThrValLeuAspThrIleSerGlnPheSerLeuTyraPile 326
DB 1864 CCGTCAAGAAATGACTTAATTTGATTTGATCTTTCGCAACTTTCTTTTATGATATA 1923
QY 327 LysArgTyArgAspSerIleGlyGlyIleGluValLysGlyIleLysAsnGluLeuThr 346
DB 1924 CGTCGTTT-----CCAAAGAGAGTAGAATCAAGAAATTAATTAAC 1959
QY 347 ArgGluIleTyThrThrGluIleAsnPheAspArgLeuProGlnLeuArgValGlnPro 366
DB 1960 AAGAGATTATTAACAAGTTAGATCAATTAACACGACACCA----- 2001
QY 367 AsnLeuAlaThrMetGluTyraAsnLeuThrArgAlaSerPheLysLeuPheSerPheLeu 386
DB 2002 -----GGGCTATTACTTGGCTG 2019
QY 387 GluGlnPheIlePheTyThrGluAsnThrAsnPheGlyAsnArgLeuValGlyIleSer 406
DB 2020 TCAGATATTGAGTTATACACGAGAGAGTGGCAGAAAGGCGCATTAATTAATCAAGTATT--- 2076

QY 407 AsnArgAspAlaProThrTySerAsnThr-----IleThrGluThrLeuTyArgly 423
DB 2077 ---CGAGAGCTAAATATATATATCTGATCAATTTTTTACAGTAAATAATTTATGCT 2133
QY 424 GluArgThrGly---SerProThrTyThrIleArgProPheGlnSerTyLysVal 442
DB 2134 AATTAACAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2190
QY 443 SerIleValThrAspArgGlnSerProProValSerProIleGlnProHisPheIleIle 462
DB 2191 CACTTAAGCATAAACCGCTCTTTCAACAATAGCTGGTAAATTAATTAATTAATTAATTA 2250
QY 463 AsnGlnIleGluLeuTyLysAsnGlySerSerAsnAsnThrLeuLysTySerAlaGly 482
DB 2251 ATTCAAAAAATGTTTTCACACTTTTAAACGATTAI-----GAATATCAAAAAAT 2304
QY 483 GlySerLeuSerAsnTy-----GlnAsnThrThrPhePheGlnPheProArgLys 499
DB 2305 TTTAATGTGATATTAACAATAATGAACCTCAAGAACTACA----- 2343
QY 500 LysAspCysAsnLeuValIleAspProGlyCysSerProAsnPheAsnAsnTySerHis 519
DB 2344 -----AACTATCTTAATATATATGATGCTTCA---AACAGCCAAAATTCAAAACAT 2391
QY 520 IleLeuSerHisPheSerLeuPheThrTySerTyValIleGlyLeuGlnLeuGlnIle 539
DB 2392 AATTATATCTCATTTTCCATTATCATCCAC-----AAGTTAGAGTTT 2433
QY 540 LeuAsp-----ThrGlyValLeuGlyTyTPThrHisSerSerValAspArgTy 555
DB 2434 GCTGAGTATTTTCACTCTATATTTGCAATGATGATGAGACACAAATGTGTAACCTCCA 2493
QY 556 AsnAlaIleSerAspLysIleIleThrMetIleProAlaIleLysGlyAsnLeuAsp 575
DB 2494 AATTAAATTAATGAGAAAGTGTGATGACAAATCCCATTTGTAACCTTACGAAGTT--- 2550
QY 576 ThrAsnSerValIleGluGlyProGlyHisThrGlyGlyAsnLeuValTyLysGln 595
DB 2551 ACTTAACATTCAGTTTATAGAGACCGAGTTTTCAGTGGAGATTAATTAATTAATTAATTA 2610
QY 596 SerGlnGlyArgLeuGluIleThrCysGluThrProAsnSerThrGlnSerTyPheIle 615
DB 2611 GAT-----AATGTCTCTAATTAATGTAAGCT---AGTTCTTAAATAAATACGCTATA 2661
QY 616 ArgLeuArgTyAlaThrAsnGlyAlaGlyAsnThrLeuProAsnIleSerLeuThrIle 635
DB 2662 AGCTATATTTATGCTGCAAAATTAACGAAATAGCTGATCAATAGACTAGCTGATTCGGA 2721
QY 636 ProGlyValIleGlyIleProProGlnArgLeuAsnAsnThrPheSerGlyThrAsnTy 655
DB 2722 GCGAGATTTCTA-----TTGCAACTTACCTTTTCTAGAAAAGGGAAC 2763
QY 656 AsnAsn-----LeuGlnTyArgLysAspPheGlyTyPheGlnPheProSer 670
DB 2764 AATTAATTTTACATTCAGACCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2823
QY 671 ThrValThrLeuProLeuAsnArgAsnIleProPheIlePheAsnArgAlaAsp----- 688
DB 2824 GATATTAATTAATCCCGAAGATGAAGAAATTAATTAATTAATTAATTAATTAATTAATTA 2883
QY 689 ValSerAsnSerIleLeuIleIleAspLysIleGluPheIleProIleThrSerSerMet 708
DB 2884 GAGAGGAGATATTTTATTAATGATTAATTAATTAATTAATTAATTAATTAATTAATTA 2943
QY 709 HisGlnAsnArgGluLysGlnLysLeuGluThrIleGlnThrLysIleAsnThrPhePhe 728
DB 2944 -----ACTAATGAATGAATTTAGAGAAAGCAAGAAAGCAATGATATTAATTT 2994
QY 729 ThrAsnHisThrLysThrLeu 735
DB 2995 ATTAACGCAAAACGCTTTG 3015

Search completed: January 20, 2006, 17:46:31
Job time : 1177 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - nucleic search, using frame plus p2n model

Run on: January 20, 2006, 16:31:25 ; Search time 9649 Seconds
(without alignments)
4329.976 Million cell updates/sec

Title: US-10-783-417-2
Perfect score: 3869
Sequence: 1 MNQNNNDNNEYEIDSHSTSPY.....KLETIQTKNPFTNHTKTL 735

Scoring table:
BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 5883141 seqs, 28421725653 residues
Total number of hits satisfying chosen parameters: 11766282

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODE=frame+ p2n.model -DEV=x1h
-O=/cgn2.1/USPTO.spool/US10783417/runat_20012006_095010_23867/app_query.fasta_1.903
-DB=GenEmbl -QFMT=fastap -SUFFIX=p2n.rge -MINMATCH=0.1 -LOOPCL=0 -LOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blomsun62 -TRANS=human40.cdi -LIST=45
-DOCLIGN=200 -THR_SCORE=DCT -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCL
-OUTFMT=pco -NOM=ext -HEAPSIZE=500 -MILEN=0 -MAXLEN=200000000
-USER=US10783417@cgn2.1 6416 @runat_20012006_095010_23867 -NCPU=6 -ICPU=3
-NO_WMAP -NRC SCORES=0 -WAIT -DSPBLOCK=100 -DEV TIMEOUT=120 -WARN TIMEOUT=30
-THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7 -YGAPOP=10
-YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl:.*
1: gb_ba:.*
2: gb_in:.*
3: gb_env:.*
4: gb_ov:.*
5: gb_ov:.*
6: gb_pac:.*
7: gb_ph:.*
8: gb_pr:.*
9: gb_ro:.*
10: gb_sts:.*
11: gb_sy:.*
12: gb_un:.*
13: gb_vl:.*
14: gb_hcg:.*
15: gb_pl:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3869	100.0	2208	6	CQ868305 Sequence
2	3314	85.7	2235	6	CQ868307 Sequence
3	3104.5	80.2	2085	6	CQ868309 Sequence

Result No.	Score	Query Match	Length	DB ID	Description
4	1061.5	27.4	3543	1	BTTOXD1
5	1061.5	27.4	4451	6	I08083
6	1061.5	27.4	4451	6	I09103
7	1060.5	27.4	4253	1	BACISRH4
8	1060.5	27.4	4934	6	E01676
9	1060.5	27.4	127923	1	BTBPTOXIS
10	897.5	23.2	3753	1	BACMSQB
11	897.5	23.2	127923	1	BTBPTOXIS
12	880	22.7	3756	6	E00614
13	862.5	22.3	4959	1	AP285775
14	850	22.0	6009	1	BTB251977
15	849	21.9	4896	1	AF132928
16	808.5	20.9	3752	1	AB125059
17	784	20.3	3644	1	BTB251978
18	775.5	20.0	3507	1	BTB04365
19	775.5	20.0	3507	6	I25972
20	759	19.6	3684	1	BTTOXD2
21	759	19.6	3684	6	I08884
22	757.5	19.6	3535	6	I06096
23	754	19.5	3684	1	BTTOX
24	752.5	19.4	4186	1	BACISRH3
25	752.5	19.4	4186	6	E01905
26	746	19.3	3642	1	AB193814
27	709	18.3	3668	1	D88381
28	705	18.2	4391	1	BTBEG65
29	705	18.2	6698	1	AB116651
30	703	18.2	3633	6	CS130948
31	703	18.2	3633	6	AK543926
32	703	18.2	6613	6	CS130963
33	703	18.2	6613	6	AK543951
34	696.5	18.0	4056	1	BACCRXD2
35	696	18.0	3536	6	E01029
36	689.5	17.8	4188	6	BD177494
37	689.5	17.8	4188	6	AK670919
38	687	17.8	3931	1	AB161456
39	678.5	17.5	2003	6	AK543930
40	678.5	17.5	2022	6	CS130958
41	678.5	17.5	2022	6	AK543940
42	677	17.5	8572	1	AB116649
43	677	17.5	8605	1	AB116650
44	676	17.5	1953	6	I25973
45	676	17.5	1977	1	BTCADE

ALIGNMENTS

RESULT 1
CQ868305
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES
source
CDS

2208 bp DNA linear PAT 13-SEP-2004
Sequence 6 from Patent WO2004074462.
CQ868305.1 GI:51998351
Bacillus thuringiensis
Bacillus thuringiensis
Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus; Bacillus cereus group.

1
Carozzi, N., Hargiss, T., Kozief, M.G., Duck, N.B. and Carr, B.
Delta-endotoxin genes and methods for their use
Patent: WO 2004074462-A 6 02-SEP-2004;
Athenix Corporation (US)
Location/Qualifiers
1..2208
/organism="Bacillus thuringiensis"
/mol_type="unassigned DNA"
/db_xref="taxon:1428"
1..2208
/note="unnamed protein product"
/codon_start=1
/transl_table=11
/protein_id="CAH33948.1"
/db_xref="GI:51998352"

|||||
Db 1801 GAAATTCAGTGAAGAACTCTAATCTACACAACTTATTCTAGACTGCATGCT 1860
QY 621 ThrAenGlyValAGlySerThrLeuProAenIleSerLeuThrIleProGlyValIleGly 640
Db 1861 ACAAATGGTGGTGAATACTCTCTCAATATATCTCTTACAAATACAGAGAAATGGA 1920
QY 641 IleProGlnIleArgLeuAenAenThrPheSerGlyThrAenTyraAenLeuGlnTy 660
Db 1921 ATRACCTTCACAGACTCAACACACTTTTTCGTGACAAATTAATATATTTACATAC 1980
QY 661 GlyAspPheGlyTyrrPheGlnPheProSerThrValThrLeuProLeuAenArgaenIle 680
Db 1981 GGAGATTTTGGGATTTCCAAATTTCCAGACAGTACATTAACCTTAATGCAACATTA 2040
QY 681 ProPheIlePheAenAenArgAlaAspValSerAenSerIleLeuIleIleAspIleGly 700
Db 2041 CCAATTATATTAATATGTCAGATGATCAAAATTCATTTATATCAATTAATAATTTGAA 2100
QY 701 PheIleProIleThrSerSerMetHISGlnAenArgGluLeuGlnLeuGluThrIle 720
Db 2101 TTTATACAAATTAATCTCTCTAATGACCAAAATAGGAAATAACAAATTAAGAACTATC 2160
QY 721 GlnThrIleValAenThrPhePheThrAenHISThrIleThrLeu 735
Db 2161 CAAACAAATAATATACATTTTTCACAAATCATACAAACAACTT 2205

RESULT 2
LOCUS CQ868307 2235 bp DNA linear PAT 13-SEP-2004
DEFINITION Sequence 8 from Patent WO2004074462.
ACCESSION CQ868307
VERSION CQ868307.1 GI:51998353
KEYWORDS

SOURCE Bacillus thuringiensis
ORGANISM Bacillus thuringiensis
Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus; Bacillus cereus group.

REFERENCE 1
AUTHORS Carozzi, N., Hargiss, T., Koziel, M.G., Duck, N.B. and Carr, B.
TITLE Delta-endotoxin genes and methods for their use
JOURNAL Patent: WO 2004074462-A 8 02-SEP-2004;
Athenix Corporation (US)

FEATURES
source Location/Qualifiers

CDS

1..2235 /organism="Bacillus thuringiensis"
/mol_type="unassigned DNA"
/db_xref="taxon:1428"
1..2235 /note="unnamed protein product"
/codon_start=1
/transl_table=11
/protein_id="CAH33949.1"
/db_xref="GI:51998354"
/translation="MNQNNNVEEIIIDSKNLVSPSNRNIDHSRYPTNNPNQNTN
YKEMLMCGANTQYGDNEFTPSADTIAVSAKTYSGTLGIGGLNISIEPIGIG
AIIISGTLITVFWPAGSDQKVTWQIFIMGSI FVDTPLESIKQIKDQLEGFQIL
QSYNTLADWRKIKRLQAPGLPSSALQAPALTCLKIRPNVNDPIREIPGQLETK
TLILPIYAQANFHLNLQCGABLADEMNADILPSQIBENAGTSDIYLLKENIPKY
SNVCANTYREGAKLNEBNPMWMSIPNDPRYWTITVLDITIAQFSYDIKRYKDSIGR
IGIKTELIREIYETREINPDLTYLEIOPNLAIMENYLRSGLRLESPIDELIEYTKN
EYGNRLVGIANRNSTATTGTEIIEYGRTPPTTKTILIPESYKVSIVYTRQVPT
SPENPTSHIINQILELYLANSPSNKGLVYSGNLSNDKTTDQFQPKCKPIINENC
LPSYNSYSHIISQLEFNSYKIGLANLITGLGWTSSVNRNNAISDKITMIPIA
IKGNSLDTNSKVLIEGHTGNNLVYLOSQGLEITCRTPNSQSYIIRLAVTNGN
TLPNTSLATPGVIGIPORLNTFSGTNNNOYDGFGEFQPSVTTLPLNNIPFI
NEADVNSLIILIDKIEFIPITSSVNRKQKLEFIIQTINIFPNTHTKNTLNTAATN
YIID"

ORIGIN
Alignment Scores:
Pred. No.:
Score:

5..39e-241 Length: 2235
3314.00 Matches: 636

Percent Similarity: 91.34
Best Local Similarity: 86.94
Query Match: 85.74
DB: 6 Gaps: 3

US-10-783-417-2 (1-735) x CQ868307 (1-2235)

QY 1 MetAenGlnAenAenAenAenAenGluTyrrGluIleIleAenSerHISThrSerProTy 20
Db 1 GCGAATCAAAATTAAT---AATTAATGAATATGATGATTAATGATTCAAAGATTAATCTTAT 57
QY 21 PheProAenAenArgAenSerAenAenSerAenArgTyrrProTyrrThrAenAenProAenGlnPro 40
Db 58 CCTTCACAGAAATATGATGATCATTCATGATACCTTACACAAATATCAATCAACCA 117
QY 41 LeuGlnAenThrAenTyrrGlyLeuThrLeuAenMetCysGlnGlyAenThrGlnTyrrGly 60
Db 118 TTAACAAACACAAATTAACAAAGAGTGGCTCAATATGTCAGAGGAAATACAAATATGCT 177
QY 61 AspAenPheGluThrPheAlaSerAlaAspThrIleAlaIleValSerAlaGlyThrIle 80
Db 178 GATTAATTCAGACATTTGCTATGCTGATACAAATTCGTCGATTAATCGGACAGTACTAT 237
QY 81 ValSerGlyThrLeuLeuAlaGlyIleGlyLeuThrSerIleSerGlyProIleGly 100
Db 238 GATCCGGTACTCTGTAGCCGTATAGTGGCTCACTTATATCCGACCGAATGAGA 297
QY 101 IleIleGlyAlaIleIleIleSerPheGlyThrLeuIleThrValPheTrpProAlaGly 120
Db 298 ATTAATAGGTGCTATATATATATCTTTGGTACCTTAATCACTGCTTTTGGCCCGGGA 357
QY 121 GluGlnAenPheThrValTrpThrGlnPheIleIleMetGluIlePheValAspThr 140
Db 358 GACACAGACAAACAGATATGACACAAATTTATTAATAGGGAATTTTGTGATACA 417
QY 141 ProLeuThrGluSerIleIleGlnLeuLeuGlnThrLeuGlnGlyPheArgGlnIle 160
Db 418 CGTTTACAGAAAGCATTAACAGCTTAAGTTACAACTTTGAAAGATTTAGCAAAATA 477
QY 161 LeuGlnSerTyrrAenThrAlaLeuAenAspTrpAlaGlyIleLeuAenArgLeuGlnAlaPro 180
Db 478 TTAACAAGCTAATAAATACGATTAATGATGATGCAAAATTAATAAAGATCAACAGCTCT 537
QY 181 GlyLeuProProSerSerAlaLeuGlnGlnAlaIleAenThrLeuIleIleAenPheGly 200
Db 538 GGATTAACCAACCATCAATGACATTAACACAGTGCCTTGAATCTTAATAATAGATTTGAG 597
QY 201 AenValHISaenAenPheIleArgGluIlePheProGlyPheGlnLeuGluThrTyrrIleThr 220
Db 598 AATGTTCAACATGATTTTATTCGAGAAATACCTGCTTCCAACTTGAACTTATTAACG 657
QY 221 LeuLeuLeuProIleTyrrAlaGlnAlaIleAenPheHISLeuAenLeuGlnGlnGly 240
Db 658 CTATTACTACTATTTATGCGCACTGCTAATTTTCAATTAATTAATTAACAACAGCT 717
QY 241 AlaGluLeuAlaAspGluTrpAenAlaAspIleHISProSerGlnIleGluProAenAla 260
Db 718 GCTGAATGGCTGAAGATGAATGCAATCAATCAATCTTCAAAATGAACCTATATGCT 777
QY 261 GlyThrSerAenAenAspTyrrTyrrLeuLeuLeuGlnAenIleProIleTyrrSerAenTyrr 280
Db 778 GGAACATCGATGACTATTAATTAACCTTTTAAGAAATATACCTTAATATAGTAATCTAT 837
QY 281 CysAlaAenThrTyrrArgThrGlyLeuLeuAenLeuAenArgGluProAenMetLeuTrp 300
Db 838 TGTGCAAAATACCTATAGAGAAAGCACTAAATAAATCTTCAAGAACCTATATGAGATG 897
QY 301 SerIlePheAenAenAspTyrrArgThrGlyMetThrIleThrValLeuAenAspThrIleSerGln 320
Db 898 AGTATATTAATGATTAATGAGATTAATGATTAATGATTAATGATTAATGATTAATGATTA 957
QY 321 PheSerLeuTyrrAspIleIleAenArgTyrrArgAenSerIleGlyGlyIleGluValIleGly 340

QY 112 LeuilethValphePrProalaglygluInasPheThrValTrpThrgInphele 131
 DB 181 CTATACACTGTCTTTGGCCCGGAGAAACAAGACAAACAGTATGACACATTTAT 240
 QY 132 IyMetGlyGluIlePheValAspThrProLeuThGluSerIlelysgInleuylsleu 151
 DB 241 AAAATGGGAGAAATTTTGTGTATACACGTTAAACGAAAGCATAAACGCTAAAGTAA 300
 QY 152 GlnThleuGluGlyPheArgGlnIleleuGlnSerTyAsnThrIleleuAspPheTP 171
 DB 301 CAAACTTAAAGAGATTAGACAAATATTAACAAAGCTATTAATACACATTAGATGATG 360
 QY 172 ArglyLeuLeuLeuArgleuGlnAlaProGlyLeuProProSerSerAlaLeuGlnAla 191
 DB 361 AGAAATTTAAAGAACTACAGCTCTGATTAACCATCATCAGCATTTACACAGCT 420
 QY 192 AlaLeuThleuLeuValleArgPheGluAsnValHisAsnAspPheIleArgGluIlePro 211
 DB 421 GCCTGACTCTTAAATACGATTGAGAAATGTTCAACATGATTTTATTCGAGAAATACCT 480
 QY 212 GlyPheGluLeuGluThrTyrlserThrleuLeuProIleTyrlalGlnAlaAlaAsn 231
 DB 481 GGTTTCACACTTAAAGCTTATTAACGCTATTAATTAATTTATGCGCAGCTGCTAAT 540
 QY 232 PheHisLeuAsnleuLeuGlnGlnGlyAlaGluLeuAlaAspGluTrpAsnAlaAspIle 251
 DB 541 TTTCAATTTAAATTTATTTACACAGAGCTGCTGAATTTGGCTGATGAAAGCATGATTA 600
 QY 252 HisProSerGlnIleGluProAsnAlaGlyTyrlserAspAspTyrlTyrlsleuLeu 271
 DB 601 CATCTTCAACAAATTTACATTAATGCTGAAACATGACATGACTATTAATTAATTTAA 660
 QY 272 GluAsnIleProLeuTyrlserAsnTyrlTyrlalAsnThrTyrlArgThrGlyLeu 291
 DB 661 GAAAAATATCTTAAATATGATTAATGATGCTGCTGCTGCTGCTGCTGCTGCTGCT 720
 QY 292 LeuArgAspGluProAsnMetLysTrpSerIlePheAsnAspTyrlArgTrpMetThr 311
 DB 721 CTTCGAAACGAACTTATTAATGATGATTAATTAATTAATTAATTAATTAATTAAT 780
 QY 312 IleThrValleuAspThrIleSerGlnPheSerleuTyrlAspIleLysArgTyrlArgAsp 331
 DB 781 ATTACTGATTAATGATTAATGATTAATTTCTTTTATGATTAATTAATTAATTAAT 840
 QY 332 SerIleGlyGlyIleGluVallysgIlyIleLysAsnGluLeuThArgGluIleTyrlThr 351
 DB 841 TCATATGGA-----AGAAATAGTGGCATTAATAACGAACTTACAGAAATTTTATTA 894
 QY 352 ThrGluIleAsnPheAspArgleuProGlnleuArgValGlnProAsnleuAlaThrMet 371
 DB 895 ACTGAAATTAATTTTACCGTCTTACTTAATCTTGAATTAATTAATTAATTAATTAAT 954
 QY 372 GluTyrlAsnleuThArgAlaSerPheLysleuPheSerPheleuGlnPheIlePhe 391
 DB 955 GAATATTAATTTTACAGTTCAGGCTTATTAATTTTCAATTTTATTAATTAATTAAT 1014
 QY 392 TyrlThrGluAsnThrAsnPheGlyAsnArgleuValGlyIleSerAsnArgAspAlaPro 411
 DB 1015 TATTAACAAAATGAAGCACTGAGGAAATCGTTTATGTTATGCGAATCGTATTAATCAT 1074
 QY 412 ThrTyrlserAsnThrIleThrGluThrleuTyrlGlyGluArgThrGlySerProThrThr 431
 DB 1075 ACTTAATGCTACGACAGAACTGAATTAATTAATTAATTAATTAATTAATTAATTAAT 1134
 QY 432 TyrlThrIleArgProPheGluSerTyrlsValSerIleValThrAspArgGlnSerPro 451
 DB 1135 AAAACTTAAATACATTTGATCTTAAATTAATTAATTAATTAATTAATTAATTAAT 1194
 QY 452 ProValSerProIle--GlnProHisPheIleIleAsnGlnIleGluLeuTyrlleuAsn 470
 DB 1195 CTTACTTCCCTTTTCTTAACATATTAATTAATTAATTAATTAATTAATTAATTAAT 1254
 QY 471 GlySerSerAsnAsnThrleuTyrlserAlaGlyGlySerleuSerAsnTyrlGlnAsn 490

DB 1255 AATTCACCTAGTAATAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1314
 QY 491 ThrThrPhePheGlnPheProArglylyAspAspCysAsnleuValIleAspProGlyCys 510
 DB 1315 ACNACTGATTTTCAATTTTCTGTATAAAAAAGCTATAAACCAATTAATTAATTAAT 1374
 QY 511 SerProAsnPheAsnAsnTyrlserHisIleleuSerHisPheSerleuPheThrTyrlSer 530
 DB 1375 TTACCAAGCTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1434
 QY 531 TyrlValIleGlyLeuGlnleuGlnIleleuAspThrGlyValleuGlyTyrlThrHisSer 550
 DB 1435 TATAAATTTGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1494
 QY 551 SerValAspArgTyrlAsnAlaIleSerAspLysIleIleThrMetIleProAlaIleLys 570
 DB 1495 AGCTTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1554
 QY 571 GlyAsnAsnleuAspThrAsnSerLysValIleGlyGlyProGlyHisThrGlyLysAsn 590
 DB 1555 GGTAAACAGCTTGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1614
 QY 591 LeuValIleGlyLeuGlnSerGlnGlyArgleuGluIleThrCysGluThrProAsnSerThr 610
 DB 1615 TTGGTTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1674
 QY 611 GlnSerTyrlPheIleArgleuArgTyrlalThrAsnGlyAlaGlyAsnThrleuProAsn 630
 DB 1675 CATCTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1734
 QY 631 IleSerleuThrlleProGlyValIleGlyIleProProGlnArgleuAsnAsnThrPhe 650
 DB 1735 ATATCTTACATTAACGAGATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1794
 QY 651 SerGlyThrAsnTyrlAsnAsnleuGlnTyrlGlyAspPheGlyTyrlPheGlnPheProSer 670
 DB 1795 TCTGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1854
 QY 671 ThrValIleuProleuAsnArgAsnIleProPheIlePheAsnArgAlaAspValSer 690
 DB 1855 ACAGTAACATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1914
 QY 691 AsnSerIleleuIleIleAspLysIleGluPheIleProIleThrSerSerMetHisGln 710
 DB 1915 AATTAATTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1974
 QY 711 AsnArgGluLysGlnleuGlnleuThrlleGlnThrLysIleAsnThrPhePheThrAsn 730
 DB 1975 AATAGGAAAAACAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2034
 QY 731 HisThrLys 733
 DB 2035 CATACAAA 2043
 RESULT 4
 BTTOX01 3543 bp DNA linear BCT 18-APR-2005
 LOCUS Bacillus thuringiensis gene for 130 kDa delta-endotoxin.
 DEFINITION Y00423
 ACCESSION Y00423.1 GI:40351
 VERSION delta-endotoxin; endotoxin.
 KEYWORDS Bacillus thuringiensis
 SOURCE Bacillus thuringiensis
 ORGANISM Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus; Bacillus cereus group.
 REFERENCE 1 (bases 1 to 3543)
 AUTHORS Ward, R.S. and Bilar, D.J.
 TITLE Nucleotide sequence of a Bacillus thuringiensis var. israelensis gene encoding a 130 kDa delta-endotoxin
 JOURNAL Nucleic Acids Res. 15 (17), 7195 (1987)
 PUBMED 2821500
 REFERENCE 2 (bases 1 to 3543)

AUTHORS
TITLE
JOURNAL

Ward, E.S.
Direct Submission
Submitted (11-MAY-1987) Ward E.S., University of Cambridge,
Department of Biochemistry, University of Cambridge, Tennis Court
Rd., Cambridge CB2 1QM
*strain: var:israelensis;
Data kindly reviewed (12-JUN-1988) by Ward E.S.
Location/Qualifiers

FEATURES
source

CDS

1..3543
/organism="Bacillus thuringiensis"
/mol_type="genomic DNA"
/db_xref="taxon:1428"
1..3543
/note="unnamed protein product; endotoxin (AA 1-1180)"
/codon_start=1
/transl_table=11
/protein_id="CAA68485.1"
/db_xref="GI:40352"
/db_xref="GOA:P16480"
/db_xref="UniProt/Swiss-Prot:P16480"
/translation="MNPYONKNEYETLNASQKLNISNNYTRPIENSPOKLOSTNY
KDLNMCQNOQYGGDFETIDSGELSAVTVLFGFTPLGLALIGFTLIP
VLFPADQSNMSEITOTKNTIKKEIASYISNANKILNLSFNVISYHNLKTMEN
NPNQNTDVRTOQLVYHQNVIPELVNCGPSPSCDYNNILVLSYQANMLH
TVLNDVAFKAKNRPDIPLPIIDYPTVTKAIEDTNTCVTVKGLMLIK
TTPSNLGNINMNTYNTRTKMTAVLDVALFPNVGKYPGIVGSELREIYVL
NFESPKYDYFOYXEDSLTRPHLFTLDSLNEYKQKTPNNEFTSHMFFYLLD
NISQSVFGNHNVTDLKSLGLATNVIPLNLNKNLNDYNNISKDDPILNG
TRLEKELIAGSGOITVDVNKRIEGLPIIKRENGNGLTPEYDNNSHLISFISLS
IPATKTOVYTPAATHSSVDPRNTITHTLTQIPAVKANSGLTASKVQGGHCGDL
IDPKHFKITCOHSPFOQSYFRIRIKASNGSNTAVNLISI PGVAILGLMLNPFSG
TDYTNLKADFOYLEFSNEVFAPQNI SLVFNRSDYVNTVLIDKLEPIRISIR
EDREKOKLETVOQIINFEYANPIKNTLSDITDIDDAALVEGISSEELPKKEMLD
LDEVNAKOLTSRNVLQNGDEBSATLGMTSDNTTIOEDPIFKGHYLSKEMLD
GTFPFTYFOKIDESKLPYTRIVYRGVGSKQVEIVSRVGEEDLIMVPADLNY
LYSPFDEGSRCTSAVPANIGTSMKSCQVDRKRVKVMDSQSPFTIDTGA
LDINENIYVWAFKISSPDGASLDNLNVLISBPIDGALSRRKMEKMGDOEAKR
SETQQAIVAAQALDALFTNVODEALQVYAEIVVOSIPYVNNMLSDVPG
MNYDIYVELIDARVAQALYDI RNIIKNGDFOGMGHNVGNADVOQIDVSLVLS
NMSAGVSONVHLQHNHGYLVGLVIAKKEGPGNGYVTLMPMEENQELTFTSDEGYITK
TVDVPPDTRVRIEIGETGSEFYIESILICMNE"
919
old_sequence
/note="g was a in [1]"
/citation=[2]
922
old_sequence
/note="g was a in [1]"
/citation=[2]
ORIGIN

Alignment Scores:

Pred. No.: 2,578-70 Length: 3543
Score: 1061.50 Matches: 271
Percent Similarity: 50.4% Conservatve: 118
Best Local Similarity: 35.1% Mismatches: 278
Query Match: 27.4% Indels: 105
DB: 1 Gaps: 24

US-10-783-417-2 (1-735) x BTOXDI (1-3543)

QY 1 MetAaNgInaenAaSpaenAaNgIuYrGluIleIleAaPSeRHiThSeRProYr 20
DB 1 ATGAATCCCTTAACAAATAAATAATGATAATGATAATGCTTCACAAAAAAATTA 60
QY 21 PheProAaNgAaSeRSeAaPSeRqTYrProYrThraAaNgProAaNgInPro 40
DB 61 -----AATATATCTAAATATATATACAGATATCCAAATAGAAATAGTCCAAACATTA 114
QY 41 LeuGInaenThraAaNgIuYrGluIleAaPSeRqTYrGluIleAaPSeRqTYrGlu 60
DB 115 TTACAAAGTAAATTAATAAGATTTGGCTCATATGTGTCAACAGAAATCAGCAATAGT 174
QY 61 AaPaenPheGluThraAaNgIuYrGluIleAaPSeRqTYrGluIleAaPSeRqTYr 80
DB ::

DB 175 GGAAGATTTGAACCTTTATTGATAGT-----GGTGAACCTGACGCTTAACTAAT 225
QY 81 ValSerGIYThrLeuLeuAaGIYIleGIYGIYLeuThraSerIleSerGIYProIleGIY 100
DB GTAGTTGGGACCGTACGATGCTGTTTGGGTTCACACA-----CCCTTAGA 273
QY 101 IleIleGIYIleIleIleIleSerPheGIYThrLeuIleThraValPheTrProAaGIY 120
DB 274 CTT-----GCTTAATAGTTTGGTTCATTAATTAACGATCTTTTCCAGCCCA 324
QY 121 GluGInaenPheThraValIleThraGluPheIleIleSerGIYIlePheValAaPheR 140
DB 325 GACCAATCTTAACACA---TGGATGACCTTATTAACACAACTTAATAATATTAAAAA 381
QY 141 ProLeuThraGluSerIleIleGIYLeuIleGIYLeuGluGIYLeuPheArgGluIle 160
DB 382 GAATATGATCAACATATATATAGTAATGCTAATAAATTTTAAACAGGTGCTTAATGT 441
QY 161 LeuGInaenThraAaNgIuYrGluIleAaPSeRqTYrPheArgIleAaPSeRqTYr 180
DB 442 ATCAGCACTTATCAATATCACTTAAACATGGAG-----AATTAATCCA 486
QY 181 GIYLeuProPheSerSerPheIleGIYIleGluIleAaPSeRqTYrPheArgGluIle 200
DB 487 AACCCAAATATCTCAGATGTAAAGACAAATCCAGCTAGTTCATTACCATTTTCA 546
QY 201 AaNgAaIleAaenAaPheIleArgGluIlePro-----GIYPheGluIleGluThra 218
DB 547 AATGTCAATCCAGACGCTTGAACCTCTGCTCTCTCTAATCTTAAGATGGATGATCAAT 606
QY 219 LysThraLeuLeuProIleThraValAaGIYIleAaPSeRqTYrPheArgGluIle 238
DB 607 AACATCTAGATTAATATCACTTAATGACCAACACAAACTTAATCACTGATTAAT 666
QY 239 GluGInaenAaGIYLeuAaPSeRqTYrPheArgGluIleAaPSeRqTYrPheArgGlu 258
DB 667 CAAGCCCTCAATTTGAAGCTAATTTAAAAAACAATGCAATTCGATTATTAGAGCT 726
QY 259 AaNgAaGIYThraSerAaPSeRqTYrThraValAaGIYIleAaPSeRqTYrPheArg 278
DB 727 ---TTGCCAAGACGATGATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 783
QY 279 AaNgIYCyAaIleAaThraIleThraValAaGIYIleAaPSeRqTYrPheArgGlu 296
DB 784 AATATATGTGTAACACTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 843
QY 297 -----AaenIleYTrSerIlePheAaPSeRqTYrPheArgGluIle 312
DB 844 AATCTGATGAAATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 903
QY 313 ThraValAaPheThraIleSerGluPheSerLeuThraValAaGIYIleAaPSeRqTYr 332
DB 904 GCTGATTAATAGTGTGTTGCACTCTTCTTAATTAATTAATTAATTAATTAATTAAT 954
QY 333 IleGIYGIYIleGluValAaGIYIleAaPSeRqTYrPheArgGluIleThraVal 352
DB 955 -----CCAAATAGTGTGCAACTTAATTAATTAATTAATTAATTAATTAATTAAT 999
QY 353 GluIleAaPheAaPheArgLeuProGluIleAaGIYIleAaPSeRqTYrPheArgGlu 372
DB 1000 ---CTTAATCTGAAAGAAAGCCCTTAATTAATTAATTAATTAATTAATTAATTAAT 1050
QY 373 ThraValAaPheThraValAaPheArgLeuPheSerPheArgGluIlePheArg 392
DB 1051 GATTCATCTTAACGATTAACG---CATTTATTACTTGCTGATTCCTTAATTTTAT 1107
QY 393 -----ThraValAaPheThraValAaPheArgLeuPheArgGluIle 401
DB 1108 GAAAGAGGCAACATCTCTTAATTAATTTTTCACAGCCATTAATTAATTAATTAATTAAT 1167
QY 402 ---LeuValGIYIleSerAaPheAaPheArgLeuPheArgGluIleThraVal 420
DB 1168 ACATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1227

QY 421 LeuTYGlyluarGthrGlySerProThrThyLeuThylleuArgProPheGluSerTYr 440
1228 TTR-----AAATCTGTGTTGGCAACAAATTTATTT----- 1263
QY 441 LysValSerlleValThrAspArgInsProValSerProIleGlnProHisPhe 460
1264 -----TTT 1266
QY 461 llelleaGlnInleGluLeu-----TyrleuGlnGlySerSerAsnThr---- 476
1267 TTAATTAATGTCATAGCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1326
QY 477 -----LeuYSerSerIleGly 483
1327 ATGATTTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1386
QY 484 SerleuSerAsnTYrGln--AsnThrThrPhePheGlnPheProArglyAspCys 502
1387 GGGCAATAACTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1446
QY 503 AsnleuValIleAspProGlyCysSerProAsnPheAsnAsnTYrSerHisIleLeuSer 522
1447 GAGAAATCAAGAAACCTTACCTTTTCCAAACATATGATTAATTAATTAATTAATTA 1506
QY 523 HisPheSerleuPheThrTYrSerTYrValIleGlyleuGlnleuGlnIleLeuAspThr 542
1507 TTTATTAATAAGTCTTATCCCTGCAACATATTAATACTCAAGTAT----- 1554
QY 543 GlyValleuGlyTYrThrHisSerSerValAspArgTYrAsnAlaIleSerAspIle 562
1555 ---ACGTTTGCTTGACACACTCTAGTGTGATCCAAAAATACATTTATACACATTTA 1611
QY 563 lleThMeIleProAlaIlelyGlyAsnAsnleuAspThrAsnSerlyValIleGly 582
1612 ACTAACCAATTCACCTGTAAAGCAATTCACCTGGGCTCTTAAGTTGTTCAA 1671
QY 583 GlyProGlyHisPheThrlyGlyAsnleuValTYrleuGlnSerGlnIlyArgleuGlyIle 602
1672 GAGCTGGTCATACAGAGGGGATTAAT-----GATTTCAAGATCATTTCAAAAT 1725
QY 603 ThrCysGluThrProAsnSerThrGlnSerTYrPheIleArgleuArgTYrAlaThrAsn 622
1726 ACATGTCACACCTCAATTTTCAACATGCTATTTTATAGATTCGTTATGCTTCAAT 1785
QY 623 GlyAlaGlyAsnThrleuProAsnIleSerleuThrIleProGlyValIleGlyIlePro 642
1786 GGAAGCGCAATATCTGAGCTGTATTAATCTTAAGTATCCAGGGGAGCAAGACTG--- 1842
QY 643 ProGlnArgleuAsnAsnThrPheSerGlyThrAsnTYrAsnAsnleuGlnTYrGlyAsp 662
1843 GGTATGCACTCAACCCACTTTTCTGTGACAGATTAATCGAATTTAAATTAAGAT 1902
QY 663 PheGlyTYrPheGlnPheProSerThrValThyleuProleuAsnArgAsnIleProPhe 682
1903 TTTCAGTACTTAAGATTTTCTTAACAGAGTCAAAATTTGCTCCAAATCAAAATCTCTT 1962
QY 683 llePheAsnArgIleAspVal---SerAsnSerIleleuIleleuAspIleGluPhe 701
1963 GGTATTAATCGTTGCGATGTATACAAACACAGACTTATTAATTAATTAATTAATTA 2022
QY 702 lleProIleTherSerSerMetHisGlnAsnArgGlyleuGlyleuGlyleuGlyIleGln 721
2023 CTGCCCAATTAATCGTTCTATAGAGAGATTAAGAGAAACAAATTAAGAAACAGTACA 2082
QY 722 ThrlyIleAsnThrPhePheThrAsnHisThrIys 733
2083 CAATTAATTAATCAATTTATGCAATCTTAATAAA 2118
Db
RESULT 5
LOCUS 108083
DEFINITION Sequence 1 from Patent EP 0296870.

ACCESSION 108083
VERSION 108083.1 GI:589204
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 4451)
AUTHORS Ellar,D.J. and Ward,E.S.
TITLE New toxin-encoding DNA fragments from Bacillus thuringiensis
JOURNAL subsp. israelensis
Patent: EP 0296870-A1 1 28-DEC-1988;
FEATURES location/Qualifiers
source 1..4451
/organism="unknown"
/mol_type="unassigned DNA"
ORIGIN
Alignment Scores:
Pred. No.: 3,42e-70 Length: 4451
Score: 1061.50 Matches: 271
Percent Similarity: 50.4% Conservative: 118
Best Local Similarity: 35.1% Mismatches: 278
Query Match: 27.4% Indels: 105
DB: Gaps: 24
US-10-783-417-2 (1-735) x 108083 (1-4451)
QY 1 MetAsnGlnAsnAsnAspAsnAsnGluTYrGlyIlelleAspSerHisThrSerProTYr 20
891 ATGAAATCCCTTATCAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 950
QY 21 PheProAsnArgAsnSerAsnAspSerArgTYrProTYrThrAsnAspProAsnGlnPro 40
951 -----AAATATTCATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1004
QY 41 LeuGlnAsnThrAsnTYrlyGlyleuTYrleuAsnMetCysGlnGlyAsnThrGlnTYrGly 60
1005 TTACAAAGACAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1064
QY 61 AspAsnPheGluThrPheAlaSerAlaAspThrIleAlaIleValSerAlaGlyThrIle 80
1065 GAGAAATTTGAACCTTTATGATAGT-----GGTGAACTCAGTCCATATCTATT 1115
QY 81 ValSerGlyThrleuLeuAlaGlyIleGlyleuThrSerIleSerGlyProIleGly 100
1116 GATGTTGGACCGTCTGACTGCTGTTTGGGTTCAACAC-----CCCTTAAGA 1163
QY 101 llelleGlyAlaIlelleSerPheGlyThrleuIleThrValPheThrProAlaGly 120
1164 CTT-----GCTTATTAATGCTTTTGGTACATTAATTAATTAATTAATTAATTAATTA 1214
QY 121 GlnGlnAspIlyThrValIlyPheThrGlnPheIlelysmecGlyIleIlePheValAspThr 140
1215 GACCAATCTTAACACA---TGAGTGACTTATTAACACAACTTAATAATTAATTAATAA 1271
QY 141 ProleuThrGlnSerIlelyGlnleuIySleuGlnThrleuGlnGlyIlyPheArgGlnIle 160
1272 GAAATAGCATCAACATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1331
QY 161 leuGlnSerTYrAsnThrAlaLeuAspAspTYrPArglySleuIyAspArgleuGlnAlaPro 180
1332 ATCAGACCTTAATCATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1376
QY 181 GlyleuProProSerSerAlaGlnGlnAlaIleAlaIleuThrleuIySleuIyAspArg 200
1377 AACCCAAATAATATCAGAGATGAAGACAAATTCAGCTAGTTCATTAACCATTTTCAA 1436
QY 201 AsnValHisAsnAspPheIleArgIlePro-----GlyPheGlnleuGlnIlyThrTYr 218
1437 AATGTCATTCAGACCTGTAAACTTGTCTCTTAATCTTAATGATTCGATTAATTAATTA 1496
QY 219 LysThrleuLeuProIleTYrAlaGlnAlaIleAsnPheIleuAsnleuGln 238

QY 41 LeuGlnAsnThrAsnTyrIysGluTrpLeuAsnMetCysGlnGlyAsnThrGlnTyrGly 60
Db 1005 TTAACAAGACAAATTAATAAGATTGGCTCAATATGTCTCAACAGAAATCAGACGATAGT 1064
QY 61 AspaenpHeGluThrPheAlaSerAlaAspThrIleAlaIleValSerAlaGlyThrIle 80
Db 1065 GGAGATTTTGAACCTTTATGTAGT-----GGTGAACCTAGCTCTATATCT 1115
QY 81 ValSerGlyThrLeuLeuAlaGlyIleGlyGlyLeuThrSerIleSerGlyProIleGly 100
Db 1116 GTAAGTGGACCGTACTGACTGCTGGTTCCGGTTCACACA-----CCCTTAGGA 1163
QY 101 IleIleGlyAlaIleIleIleSerPheGlyThrLeuIleThrValPheThrProAlaGly 120
Db 1164 CTT-----GCTTAATAGCTTTGGTACCTTAATCCAGCTCTTTCCAGCCCA 1214
QY 121 GlnGlnAspPheThrValIlePheGlnPheIleIysMetGlyGluIlePheValAspThr 140
Db 1215 GACCAATCTAACACA---TGAAGTACTTATTAACAACAATAAATATTAATAAANA 1271
QY 141 ProLeuThrGluSerIleIysGlnLeuIysLeuGlnThrLeuGlnGlyPheArgGlnIle 160
Db 1272 GAAATAGCATCAACATATATTAAGTATGCTAATAAATTTTAAACAGCTCTTAATGT 1331
QY 161 LeuGlnSerTyrAsnThrAlaLeuAspAspTrpArgIysLeuIysArgLeuGlnAlaPro 180
Db 1332 ATCAGACTTATCATATATCACTTAACAACATGGAG-----AATATCCA 1376
QY 181 GlyLeuProProSerSerAlaLeuGlnGlnAlaAlaLeuThrLeuIysIleArgPheGlu 200
Db 1377 AACCCAAATAATCACTGAGATGTAAAGACAAATCCAGCTAGTTCATTAACATTTTCA 1436
QY 201 AsnValHisAsnAspPheIleArgGluIlePro-----GlyPheGlnLeuGluThrTyr 218
Db 1437 AATGTCATTCAGAGCTGTAAACTTGTCTCTTAATCCATGATTCATTAATCATAT 1496
QY 219 LysThrLeuLeuLeuProIleTyrAlaGlnAlaAlaAsnPheHisLeuAsnLeuGln 238
Db 1497 AACATATAGTATTTATCTAGTATGACAAAGCAAACTTACATCTGATCTATTAAT 1556
QY 239 GlnGlyAlaGluLeuAlaAspGluTrpAsnAlaAspIleHisProSerGlnIleGluPro 258
Db 1557 CAAGCCGTCAAAATTTGAAGCGTATTTAAAAAACATCGCAATTCGATTAATTAAGCT 1616
QY 259 AsnAlaGlyThrSerAspAspTyrTyrIysLeuLeuIysGluAsnIleProIysTyrSer 278
Db 1617 ---TTGCCAAGCAGCAATGATTAATATTCAGATTAAGCTAAGCTATGAAGATTAACACT 1673
QY 279 AsnTyrCysAlaAsnThrTyrArgThrGlyLeuIysAsnLeuArgAspGluPro----- 296
Db 1674 AATTAATGTGTAAACAATTATTAATAAAGATTAAATTAATTAAACGACGCTGATAGT 1733
QY 297 -----AsnMetLysTrpSerIlePheAsnAspTyrArgArgTyrMetThrIle 312
Db 1734 AATCTTGATGAAATTAATTAAGTGAACATACATACGATTCGAAACAAAATGACTACT 1793
QY 313 ThrValLeuAspThrIleSerGlnPheSerLeuTyrAspIleIysArgTyrArgAspSer 332
Db 1794 GCGTATTAAGATGTTGTCACCTTCTTAATTAATGATGTAAGTAATAAT----- 1844
QY 333 IleGlyGlyIleGluValIysGlyIleIysAsnGlnLeuThrArgGluIleTyrThrThr 352
Db 1845 -----CCATAGGTGTCTCAATCTGAATCTGAACTTCGAAATTTATACAGTA 1889
QY 353 GlnIleAsnPheAspArgLeuProGlnLeuArgValGlnProAsnLeuAlaThrMetGlu 372
Db 1890 ---CTTAACCTTGAAGAAAGCCCTTAATATATAT-----GACCTTCAATATCAAGAG 1940
QY 373 TyrAsnLeuThrArgIleSerPheIysLeuPheSerPheLeuGlnGlnPheIlePheTyr 392
Db 1941 GATTAACCTTAACGTAAGACG---CATTAATTTACTGCTTGAATTTTAT 1997

QY 393 -----ThrGluAsnThrAsnPheGlyAsnArg----- 401
Db 1998 GAAAAAGCGCAACACTACCTCAATTAATTTTTCACACAGCACTTAATATATGTTCAATAC 2057
QY 402 ---LeuValGlyIleSerAsnArgAspAlaProThrTyrSerAsnThrIleThrGluThr 420
Db 2058 ACACCTGTAATATATATCCCAAAAATCTAGTGTTTTGGAAATACAAATGTAATCAATAA 2117
QY 421 LeuTyrGlyIysArgThrGlySerProThrThrLysThrIleArgProPheGluSerTyr 440
Db 2118 TTA-----AATCTCTGGTGTGGCAACAAATATTAAT----- 2153
QY 441 LysValSerIleValThrAspArgIleSerProProValSerProIleGlnProHisPhe 460
Db 2154 -----TTT 2156
QY 461 IleIleAsnGlnIleGluLeu-----TyrLeuAsnGlySerSerAsnAsnThr--- 476
Db 2157 TTATTAATATGTCATAGCTAGATTAATTAATATCTAATGATTAATATATATTAATAGTAA 2216
QY 477 -----LeuLysTyrSerAlaGly 483
Db 2217 ATGATTTTATTAATACTAATAGTATCTAGACTTTTGGAAAGAACTTACAGCAGATCT 2276
QY 484 SerLeuSerAsnTyrGln---AsnThrThrPhePheGlnPheProArgLysIysAspCys 502
Db 2277 GGGCAAAATTAATCTATGATTAATTAATAATTTTCCGGTATCAATCTTAACGAAGA 2336
QY 503 AsnLeuValIleAspProGlyCysSerProAsnPheAsnAsnTyrSerHisIleLeuSer 522
Db 2337 GAGAAATCAAGAAACCTACCTTTTCCACATATGATTAATCAATATAGTATATTAATCA 2396
QY 523 HisPheSerLeuPheThrTyrSerTyrValIleGlyLeuGlnLeuGlnIleLeuAspThr 542
Db 2397 TTTATTAAGCTTATGATTCCTGCAACATTAATAACTCAAGTAT----- 2444
QY 543 GlyValLeuGlyTyrThrHisSerSerValAspArgTyrAsnAlaIleSerAspLysIle 562
Db 2445 ---ACGTTGCTTGACACACTCTAGTGTGATCTTAATAATCAATTTATACACTTAA 2501
QY 563 IleThrMetIleProAlaIleIysGlyAsnAsnLeuAspThrAsnSerLysValIleGlu 582
Db 2502 ACTACCCAAATTTCCAGCTGTAATAAGCAATCACTTGGAGACTGCTTAAGCTTCA 2561
QY 583 GlyProGlyHisThrGlyGlyAsnLeuValTyrLeuGlnSerGlnGlyArgLeuGluIle 602
Db 2562 GGACCTGCTCATACGAGGGGATTTAAT-----GATTCAAAAGATCATTTCAAAAT 2615
QY 603 ThrCysGluThrProAsnSerThrGlnSerTyrPheIleArgLeuArgTyrAlaThrAsn 622
Db 2616 ACATGTCACACTCAAAATTTTCAACAAATCGTATTTATTAAGATTCGTATGCTTCAAT 2675
QY 623 GlyAlaGlyAsnThrLeuProAsnIleSerLeuThrIleProGlyValIleGlyIlePro 642
Db 2676 GGAAGCGCAATATCTCGAGCTGTTAATATCTTAATATCCAGGGGTGACAGACTG--- 2732
QY 643 ProGlnArgLeuAsnAsnThrPheSerGlyThrAsnTyrAsnAsnLeuGlnTyrGlyAsp 662
Db 2733 GGTATGCACTCAACCCCACTTTTCTGTTACAGATTTACGAATTTAAATATTAAGAT 2792
QY 663 PheGlyTyrPheGlnPheProSerThrValIleThrLeuProLeuAsnArgAsnIleProPhe 682
Db 2793 TTTCACTGATCTTAAGATTTTCTTAACAGAGGTGAATTTGCTCCAAATCAAAACATATCTCT 2852
QY 683 IlePheAsnArgAlaAspVal---SerAsnSerIleLeuIleIleAspLysIleGluPhe 701
Db 2853 GTGTTTATCGTTCGATGATATATACAAACACAAACAGTATTAATGAATTAATTAATTT 2912
QY 702 IleProIleThrSerSerMetHisGlnAsnArgGluIysGlnLysLeuGlnThrIleGln 721
Db 2913 CTGCCAATTAATCTCGTTCATATAGAGAGATGAGAGAAACAAATTTAGAAACAGTACAA 2972
QY 722 ThrLysIleAsnThrPhePheThrAsnHisThrLys 733


```

Db      1347 -----CCAATAGGTGTCATCTGAAGTCTGAGAAATTTATCAGGTA 1391
Qy      353 GIUIleasnPhaeapargleuPrognleuargValGlnProasnleuAlaIlethmetGlu 372
Db      1392 ---CTTAACCTCGAAGAAAGCCCTTAATAATATAT- ---GACTTCAATATCAAGAG 1442
Qy      373 TyrAsnleuThraGlnIleSerPheleuSleuPheSerPheleuGlnIlePheleuThyr 392
Db      1443 GATTGACTTACACGTAGACGCG---CATTTATTACTGCGCTGATCTTGGAATTTTAT 1499
Qy      393 -----ThrluAnthrAanpheGlyAanarg----- 401
Db      1500 GAAAAAGCGCAAACTACTCTCTAATATTTTTCACAGCATTAATATATGTTTCATAC 1559
Qy      402 ---leuValGlyIleSerAsnIleArgAlaProThrIleThySerAsnThrIleThrGluThr 420
Db      1560 ACACCTGATATATATATCCCAAAATCTAGTGTGTTTGGAAATCACAAATATACATGTAATA 1619
Qy      421 LeuTyrGlyIleArgThrGlySerProThrIleThrIleArgProPheGluSerTyr 440
Db      1620 TTA-----AAATCTTGTTGGTGGACAAATATTAATTT----- 1655
Qy      441 IysValSerIleValThrAspArgInsProProValSerProIleGlnProHisPhe 460
Db      1656 -----TTT 1658
Qy      461 IleIleAsnGlnIleGluLeu-----TyrLeuAsnGlySerSerAsnThr----- 476
Db      1659 TTATATAATAGTCATAGCTTAGATATAATAATATCTAAATGATTAATATATGATAA 1718
Qy      477 -----LeuIleTyrSerIleGly 483
Db      1719 ATGCAATTTTATATACTAATAGTACTAGACTTTTGGAGAAAGCACTTACAGAGATCT 1778
Qy      484 SerLeuSerAsnTyrGln---AsnThrIlePhePheGlnPheProArgIleIleAspCys 502
Db      1779 GGGCAATATACCTATATGATGATAATAATAATATTTGGGTTCACATCTTAAACGAAGA 1838
Qy      503 AsnLeuValIleAspProGlyCysSerProAsnPheAsnAsnTyrSerHisIleLeuSer 522
Db      1839 GAGAAATCAAGAAACCTACCTTTTCCAAACATGATGATGATGATGATGATGATGATGAT 1898
Qy      523 HisPheSerLeuPheThrIleTyrSerTyrValIleGlyLeuGlnIleLeuAspThr 542
Db      1899 TTTATATAAAGTCTTAGTATCCCTGCAACATATAAACTCAAGTGTAT----- 1946
Qy      543 GlyValleuGlyTyrThrHisSerSerValAspArgTyrAsnAlaIleSerAspIle 562
Db      1947 ---ACGTTGCTTGACACACTGATGATGATGATGATGATGATGATGATGATGATGATGAT 2003
Qy      563 IleThrMetIleProAlaIleIleGlyAsnAsnLeuAspThrAsnSerIleValIleGlu 582
Db      2004 ACTACCAATTCACAGCTGTAAGCGAATTCACCTGGAGACGCTTCTAAGGTGTTCA 2063
Qy      583 GlyProGlyHisThrGlyIleAsnLeuValIleTyrLeuGlnSerGlnIleArgLeuGluIle 602
Db      2064 GGAACCTGGTATACAGAGAGGAGATTAAAT-----GATTTCAAAATCATTTCAAAAT 2117
Qy      603 ThrCysGluThrProAsnSerThrGlnSerTyrPheIleArgLeuArgTyrAlaThrAsn 622
Db      2118 ACATGTCACACTCAAAATTTTCAACATGATGATTTTATAGAAATTCGTTATGCTTCAAT 2177
Qy      623 GlyAlaGlyAsnThrLeuProAsnIleSerLeuThrIleProGlyValIleGlyIlePro 642
Db      2178 GGAAGGCAAAATACGAGCTGTATATAATCTTAGATCCAGGGGTACAGAGATG--- 2234
Qy      643 ProGlnArgLeuAsnAsnThrPheSerArgIleThrAsnTyrAsnAsnLeuGlnIleTyrGlyAsp 662
Db      2235 GGTATGCACTCAACCCCACTTTTCTGTACAGATTAATACGAATTTAAATATAAGAT 2294
Qy      663 PheGlyTyrPheGlnPheProSerThrValIleLeuProLeuAsnArgAsnIleProPhe 682
Db      2295 TTTCAGTACTTAGAATTTTCTAAGAGGTGAATTTGCTCCAAATCAAAACATATCTCT 2354

```

```

Qy      683 IlePheAsnArgAlaAspVal---SerAsnSerIleLeuIleIleAspSerHisIleGluPhe 701
Db      2355 GGTATTAATCGTTCGATGTATATACAAACACACAGTACTTATGATTAATTTGAATTT 2414
Qy      702 IleProIleThrSerSerMetHisGlnAsnArgIleTyrGlnIleLeuGlnIleGluIleGln 721
Db      2415 CTGCCAATTAATCTCGTTCTTATTAAGAGGATAGAGGAAACAAAAATTTAGAAACAGTACAA 2474
Qy      722 ThrIleIleAsnThrPhePheThrAsnHisThrIle 733
Db      2475 CAATATATTAATATATCAATTTATGCAAAATCTATATAA 2510

RESULT 8
E01676      4934 bp      DNA      linear      PAT 29-SEP-1997
LOCUS      E01676
DEFINITION      DNA sequence encoding Bacillus thuringiensis insectivorous protein.
ACCESSION      E01676.1 GI:2169929
VERSION      JP 1988230090-A/1.
KEYWORDS      Bacillus thuringiensis
SOURCE      Bacillus thuringiensis
ORGANISM      Bacillus thuringiensis
              Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus; Bacillus
              cereus group.
REFERENCE      1 (bases 1 to 4934)
AUTHORS      Komano, T. and Himeno, M.
TITLE      INSECTICIDAL PROTEIN OF STRAIN BACILLUS THURINGIENSIS VAR
JOURNAL      Patent: JP 1988230090-A 1 26-SEP-1988;
              SUMITOMO CHEM CO LTD
COMMENT      OS      Bacillus thuringiensis
              PN      JP 1988230090-A/1
              PD      26-SEP-1988
              PF      19-MAR-1987 JP 1987066844
              PI      KOMANO TORU, HIMENO MICHIO
              PC      C12N15/00,A01N63/00,C12N1/20,C12P21/02//A01N63/02,(C12N1/20,
              CC      C12R1.19);
              CC      strandedness: Double;
              CC      topology: Linear;
              CC      hypothetical: No;
              CC      anti-sense: No;
              CC      *source: strain=HD522;
              FH      Key      Location/Qualifiers
              FT      S/UTR      621..1368
              FT      CDS      1369..4911
              FT      /product=/'Bacillus thuringiensis insectivorous
              /protein'.
              FT      Location/Qualifiers
              source      1..4934
              /organism='Bacillus thuringiensis'
              /mol_type='genomic DNA'
              /db_xref='taxon:1428'

ORIGIN
Alignment Scores:
Pred. No.:      4,63e-70      Length:      4934
Score:      1060.50      Matches:      271
Percent Similarity:      50.4%      Conservative:      118
Best Local Similarity:      35.1%      Mismatches:      278
Query Match:      27.4%      Indels:      105
DB:      6      Gaps:      24

US-10-783-417-2 (1-735) x E01676 (1-4934)
Qy      1 McrAnGlnAsnAspAsnAsnGluTyrGluIleIleAspSerHisIleThrSerProTyr 20
Db      1369 ATGAACTCTTATCAAAATATAAATATGATATGAAACATTAAGCTTCACAAAAAAATTA 1428
Qy      21 PheProAsnArgAsnSerSerAsnAspSerArgTyrProTyrThrAsnAsnProAsnGlnPro 40
Db      1429 -----AATATATCTAATATATATCAAGATATCAAGATAAGAAATATGTCCAAAACATTA 1482
Qy      41 LeuGlnAsnThrAsnTyrIleGluIleThrLeuAsnMetCysGlnGlyAsnThrGlnTyrGly 60

```

1483 TTACAAAGTACAAATATATAAGATGGCTCATATATGTGCACAGAAATCAGAGATATGCT 1542
AspaenPheGluThrPheAlaSerAlaAspThrIleAlaIleValSerAlaGlyThrIle 80
1543 GGAGATTGGAACTTTTATTCATAGT-----GGTGAACCTCAGTCCATATCTAT 1593
ValSerGlyThrLeuAlaGlyIleGlyGlyLeuThrSerIleSerGlyProIleGly 100
1594 GTAGTTGGAGCCGCTACTGCTGTTCCGGTTCACAAAC-----CCCTTAGGA 1641
IleIleGlyAlaIleIleIleSerPheGlyThrIleIleThrValPheThrProAlaGly 120
1642 CTT-----GCTTATATAGTTTGGTGTACATTAAATACAGTCTTTTCCACC 1692
GluGlnAspLeuThrValThrThrGlnPheIleIleIleIleIleValPheValAspThr 140
1693 GACCAATCTAACACA-----TGGAGTGCCTTTATTAACAACAATTTATTTAAAAA 1749
ProLeuThrGluSerIleIleGlyIleLeuIleIleIleIleIleIleIleIleIleIle 160
1750 GAATATAGCATCAATATATAGTATAGCTATATAAATTTTAAACAGGTCGTTATAGTT 1809
LeuGlnIleThrValAspThrAlaLeuAspAspThrValGlyIleValGlyAlaPro 180
1810 ATCAGACCTTATCATATCATCACCCTTAAACATGGAG-----AATATATCA 1854
GlyLeuProProSerSerAlaLeuGlnIleAlaIleIleIleIleIleIleIleIleIle 200
1855 AACCCACAAATATCTCAGATGTATAGACACAAATCCAGCTAGTTCAATTCATTTTCA 1914
AsnValIleAsnAspPheIleArgIlePro-----GlyPheGlnLeuGlnThr 218
1915 AATGTCATTCAGACCTGTATAACTCTGCTCCCTCAATCCATAGATGGCATTCATAT 1974
LysThrLeuLeuLeuProIleIleIleIleIleIleIleIleIleIleIleIleIleIle 238
1975 AACATCATAGTATATCTAGTATATCAGACAGCAAACTTACATCTCATCTATTAAT 2034
GlnGlyAlaGluLeuAlaAspGluThrPheAsnIleIleIleIleIleIleIleIleIle 258
2035 CAGCCGCTCAATTTGAAAGCGTATTTAAAAAAGCAATCCAGCAATTCGATTTATTAAGCCT 2094
AsnAlaGlyThrSerAspAspThrIleIleIleIleIleIleIleIleIleIleIleIle 278
2095 ---TTGCCAACAAGCATGATTTATTCAGATTTGACATTAAGCATTAAGCATTAACCT 2151
AsnIleIleIleIleIleIleIleIleIleIleIleIleIleIleIleIleIleIleIle 296
2152 AATATATGTGTAACTTATTAAGATTTAAATTTTAAATTTAAATTTAAATTTAAATTTAA 2211
AsnIleIleIleIleIleIleIleIleIleIleIleIleIleIleIleIleIleIleIle 297
2212 AATCTTGATGAAATATATACTGAAACATACATACATACATACATACATACATACATACAT 2271
ThrValLeuAspThrIleSerGlnPheSerLeuIleIleIleIleIleIleIleIleIle 332
2272 GGTGATATAGATCTTGTGCACTTCTTCTATATATATATATATATATATATATATATAT 2322
IleGlyGlyIleGlyAlaIleIleIleIleIleIleIleIleIleIleIleIleIleIle 352
2323 -----CCAATAGTGTCCATCTGAATCTGAATCTGAATCTGAATCTGAATCTGAAT 2367
GluIleAsnPheAspArgLeuProGlnLeuArgValGlnProAsnLeuAlaThrMetGlu 372
2368 ---CTTAACTTCGAAAGAAAGCCCTTATTAATTTAT-----GACTTCATATATCAAGG 2418
TyrAsnLeuThrArgAlaSerPheIleIleIleIleIleIleIleIleIleIleIleIle 392
2419 GATTACCTTACACGTAGACG---CATTTATTTACTGCTGATTTCTTGAATTTTAT 2475
ThrGluAsnThrAsnPheGlyAsnArg----- 401

2476 GAAAAAGCCAACTACTCTTAATATTTTTCACACAGCCATTTATATATGTTTCATTAC 2535
LeuValGlyIleSerAsnArgAspAlaProThrIleIleIleIleIleIleIleIleIle 420
2536 ACACCTGATATATATATCCAAAAATCTAGTGTGTTTGGAAATCACAAGTAACTATAA 2595
LeuIleGlyIleIleIleIleIleIleIleIleIleIleIleIleIleIleIleIleIle 440
2596 TTA-----AAATCTGTTGGCAACAATATATTTATTT----- 2631
LysValSerIleValThrAspArgIleSerProProValSerProIleGlnProIlePhe 460
2632 -----TTT 2634
IleIleAsnGlnIleGluLeu-----TyrLeuAsnGlySerSerAsnAsnThr--- 476
2635 TTATTTAATGTCATAGCTTATGATATATATATATATATATATATATATATATATAT 2694
LeuIleIleIleIleIleIleIleIleIleIleIleIleIleIleIleIleIleIleIle 483
2695 ATGATTTTTTATTAATATAGTACTAGACTTTTGGAGAAAGAACTTACAGAGATCT 2754
SerLeuSerAsnIleIleIleIleIleIleIleIleIleIleIleIleIleIleIleIle 502
2755 GGGCAATTAATCTATGATGTAATAAATATTTTGGGCTTACCAATTTCTTAAACGAAGA 2814
AsnLeuValIleAspProGlyCysSerProAsnPheAsnAsnIleIleIleIleIleIle 522
2815 GAGATCAAGAAACCCATCCCTTTTCCACATATGATATATATATATATATATATATAT 2874
HisPheSerLeuPheThrIleIleIleIleIleIleIleIleIleIleIleIleIleIle 542
2875 TTTATTAATAAGCTTATGATATCTTCGCAATATATAAATCTCAAGGTAT----- 2922
GlyValLeuGlyIleIleIleIleIleIleIleIleIleIleIleIleIleIleIleIle 562
2923 ---ACGTTGCTTGACACCTCTAGTGTATGCTTAAATATCAATTTATATACATTTA 2979
IleThrMetIleProAlaIleIleGlyIleAsnAsnLeuAspThrAsnSerIleValIleGlu 582
2980 ACTACCCAAATTCAGCTGTATTAAGCAATTCACCTGGAGCTGCTTCTTAAGTTTTCAA 3039
GlyProGlyIleIleIleIleIleIleIleIleIleIleIleIleIleIleIleIleIle 602
3040 GACCTGCTCATACAGAGGAGATTTATTT-----GATTTCAAGATCATTTCAAAAT 3093
ThrCysGluThrProAsnSerThrGlnIleIleIleIleIleIleIleIleIleIleIle 622
3094 ACATGTCAAACCTCAAAATTTTCAACAATCGTATTTATTAAGATTCGTTATGCTTCAAT 3153
GlyAlaGlyAsnThrLeuProAsnIleSerLeuThrIleProGlyValIleGlyIlePro 642
3154 GGAAGCGCAATATCTCGAGCTGTATATATCTTACTATCCAGGGGTGCGAGAACTG--- 3210
ProGlnArgLeuAsnAsnThrPheSerGlyThrAsnIleIleIleIleIleIleIleIle 662
3211 GGATGCGCATCAACCCCATTTTCTGTCAGATTTATACGAATTTAAATATATAAGAT 3270
PheGlyIleIleIleIleIleIleIleIleIleIleIleIleIleIleIleIleIleIle 682
3271 TTTCAAGTACTTAAATTTTCTTAACGAGTGAATTTGCTCCAAATCAAAACATATCTCT 3330
IlePheAsnArgAlaAspVal---SerAsnSerIleIleIleIleIleIleIleIleIle 701
3331 GTGTTTATTCGTCGATGATATATACAAACACACAGATCTTATTTGATTAATTTGATTT 3390
IleProIleIleIleIleIleIleIleIleIleIleIleIleIleIleIleIleIleIle 721
3391 CTGCCAATTTACTCGTTATATAGAGATGAGAGAAACAAATTTGAAAACAGTACAA 3450
ThrIleIleIleIleIleIleIleIleIleIleIleIleIleIleIleIleIleIleIle 733
3451 CAAATTAATTAATATATATATGCAATCTTATATAA 3486


```

/note="HMPfem hit to PF02899, Phage integrase, N-terminal SAM-like domain"
RBS complement (5283) .5286)
gene complement (5364) .5507)
/gene="pBC006"
complement (5364) .5507)
/gene="pBC006"
/note="Similar to Bacillus anthracis pXol-17 TR:O9XZY8 (EMBL:AF065404) (47 aa) fasta scores: E(): 2.1e-12, 68.08% id in 47 aa"
/codon start=1
/product="putative integral membrane protein"
/protein_id="CAD30068.1"
/db_xref="GI:21685415"
/db_xref="UniProt/TrEMBL:O8KNX2"
/translaton="MAVLKDSNYTEMNTYIEIFPKGIFMGLIFVVPFMSIATFIW LCK"
misc_feature complement (5370) .5435)
/gene="pBC006"
/note="1 probable transmembrane helix predicted for pBC006 by TMHMM2.0"
complement (5515) .5519)
complement (6451) .8160)
/gene="pBC007"
complement (6451) .8160)
/gene="pBC007"
/note="pBC007"
/note="Similar to Bacillus anthracis pXol-16 TR:O9XZY7 (EMBL:AF065404) (569 aa) fasta scores: E(): 0.96.13% id in 569 aa, and to Bacillus thuringiensis pXol orf16-like protein TR:CA050562 (EMBL:AJ296638) (310 aa) fasta scores: E(): 6.5e-122, 99.67% id in 310 aa"
/codon start=1
/transl table=11
/product="conserved hypothetical protein"
/protein_id="CAD30069.1"
/db_xref="GI:21685416"
/db_xref="UniProt/TrEMBL:O8KNX1"
/translaction="MGKSFIIQNRSLNGKTTAAVVVTLKSMQKFAVAFEEKYSYKD TYLIDILQHTVEERLADVEGFVSKAPSPSRAMSLAKTGKRLKGLITNADLEV FLVLRKCGFTNGVIVPNTYTHMMEDVQYKEEPAYIOHQSPYALAKLNLNISTEN GLDRLYTIKLTHPANESTERKANPVIYSPITFKDDFKSIANKULFDIMNQOHTET TLKSLKQDERGKNTKTFPGMTRFLAKKYPHQIRAVIEBLTALPCTGNPLFKICMQ KGVHTKRRYTLTYLISHSDFCSKEAGEEQRHPFPKPYTAKAPITAVIQENKIG ELSDNNKFIHLVLCYCHROIISVYIRGLRDMVDPRKEGYPTKLYTLKLLMDSQYQI LDTAKGEIVPLIAQHI PKERNSRDEQAVNPGIHYSMYSLNIIKMPFNVAHLKOK PAVVTEESYRNTLKQDETLPRKAYVDQGVNLAHYALIEIMREKLVAVGHKERTI PSDVREMFIEAIDTLPOEKRLVIELPQNLLEFMTKEFLVAGVTTAPDOVLHAM EIK"
RBS complement (8168) .8173)
RBS 8947) .8951
RBS 8959) .10653
gene /gene="pBC009"
8959) .10653
/gene="pBC009"
/note="Similar to Bacillus anthracis pXol-14 TR:O9XZY5 (EMBL:AF065404) (564 aa) fasta scores: E(): 1.6e-191,
Alignment Scores:
Pred. No.: 2.73e-68 Length: 127923
Score: 1060.50 Matches: 271
Percent Similarity: 50.4% Conservative: 118
Best Local Similarity: 35.1% Mismatches: 278
Query Match: 27.4% Indels: 105
DB: 1 Gaps: 24
US-10-783-417-2 (1-735) x BTPBTOXIS (1-127923)
OY 1 MetaAngliAnaAnaAspAsnGluTyrGlyIleIleApeSerHisThrSerProtyr 20
Db ATGAATCTTATCAAAATTAATAAATGAATATGAAACATTAAATGCTTCACAAAAAATTA 96465
21 PheProAnaArgAnaSerAnaSeraSeraArgTyrProTyrThiAnaSnProAnaGlnPro 40

```

[illegible]

Db 95478 GATTCACCTTACAGTAGACG---CAATTATTTACTGGCTGATTCTTTGGAATTTTAT 95422
 Qy 393 -----ThrgluAsnThrAsnPhgGlyAsnArg----- 401
 Db 95421 GAAAAAGCCAAACTACCTCTTAATATTTTTCACAGCAATTAATATGTTTCAATTAC 95362
 Qy 402 ---LeuValGlyIleSerAsnArgAspAlaProThrTySerAsnThrIleThrgluThr 420
 Db 95361 ACACCTGATATATATATCCAAAAATCTAGTGTTTTGAAATCACAACTGAATATAA 95302
 Qy 421 LeuTyGlyGluArgThrglySerProThrTySerThrIleArgProPhgGluSerTyR 440
 Db 95301 TTA-----AAATCTGTGGTTGGCAACAATATTATATTT----- 95266
 Qy 441 LysValSerIleValThrAspArgGlnSerProValSerProIleGlnProHisPhe 460
 Db 95265 -----TTT 95263
 Qy 461 IleIleAsnGlnIleGluLeu-----TyLeuAsnGlySerSerAsnAsnThr--- 476
 Db 95262 TTATTAAATGCTAAGCTTAAGTAATAATATCTAAATGATTATATATATTAGTAAA 95203
 Qy 477 -----LeuLysTySerIleArgIleGly 483
 Db 95202 ATGATTTTATTAATACTAATGCTACTAGACTTTTGAGAAAAGAACTTACACAGGATCT 95143
 Qy 484 SerLeuSerAsnTyGln---AsnThrThrPhePheGlnPheProArgIleLysAspCys 502
 Db 95142 GGGCAAAATTAATTATGATGTAATATAAAATATTTTGGGTTACCAATCTTAAACAGAGA 95083
 Qy 503 AsnLeuValIleAspProGlyCysSerProAsnAsnAsnTySerHisIleLeuSer 522
 Db 95082 GAGAAATCAAGAAACCTTACCTTTCCACATATGATGATGATGATGATGATGATGAT 95023
 Qy 523 HisPheSerLeuPheThrTySerTyValIleGlyLeuGlnIleLeuAspThr 542
 Db 95022 TTTATTAATAAGCTTAGTATCCCTGCACATATATAAACAAGTAT----- 94975
 Qy 543 GlyValLeuGlyTTrpThiSerSerValAspArgTyAsnAlaIleSerAspLysIle 562
 Db 94974 ---ACGTTGCTGGACACACTAGCTGATGATGATGATGATGATGATGATGATGAT 94918
 Qy 563 IleThrMetIleProAlaIleLysGlyAsnAsnLeuAspThrAsnSerLysValIleGlu 582
 Db 94917 ACTACCAAAATTCACCTGTAAGCGAAATTCCTGGAGCTGCTTCTTAAGGTTGTTCA 94858
 Qy 583 GlyProGlyHisThrGlyGlyAsnLeuValTyLeuGlnSerGlnGlyArgLeuGluIle 602
 Db 94857 GGAACCTGGTCATACGAGGAGGATTAATTT-----GATTTCAAGATCATTTCAAAAT 94804
 Qy 603 ThrCysGluThrProAsnSerThrglnSerTyPheIleArgLeuArgTyAlaThrAsn 622
 Db 94803 ACATGTCACACACTCAAAATTTTCAACATGATGATTTTAAAGAAATGTTATGCTTCAAT 94744
 Qy 623 GlyAlaGlyAsnThrLeuProAsnIleSerLeuThrIleProGlyValIleGlyIlePro 642
 Db 94743 GGAAGCGCAAAATTCGAGCTGTTATTAATCTTAGATCCAGGGGTAAGCAGAACTG--- 94687
 Qy 643 ProGlnArgLeuAsnAsnThrPheSerGlyThrAsnTyAsnAsnLeuGlnTyArgIleAsp 662
 Db 94686 GGTATGCGACTCAACCCCACTTTTCTGGTACAGATTATACGAATTTAAATTAAGAAT 94627
 Qy 663 PheGlyTyPheGlnPheProSerThrValThrLeuProLeuAsnArgAsnIleProPhe 682
 Db 94626 TTTCACTACTTGAATTTTCTACAGAGTGAAATTTGCTCCAATCAAAACATATCTCT 94567
 Qy 683 IlePheAsnArgAlaAspVal---SerAsnSerIleLeuIleIleAspLysIleGluPhe 701
 Db 94566 GGTATTAATGCTTGGATGATATATACAAACACAGTACTTATATTAATTAATTAATTT 94507
 Qy 702 IleProIleThrSerSerThrHisGlnAsnArgGluLysGlnIleGluThrIleGln 721
 Db 94506 CTGCCAATTAATCTGTTCTATTAAGAGAGATAGAGAAACAAAAATTTAGAAACAGTACAA 94447

Qy 722 ThrLysIleAsnThrPhePheThrAsnHisThrLys 733
 Db 94446 CAAATATTAATTAATCACTTTTATGCAAAATCTCTAATAA 94411
 RESULT 10
 BACMSOB 3753 bp DNA linear BCT 26-APR-1993
 LOCUS B.thuringiensis insecticidal endotoxin gene, complete cd.
 DEFINITION M12662
 ACCESSION M12662 GI:143228
 VERSION M12662.1
 KEYWORDS
 SOURCE Bacillus thuringiensis
 ORGANISM Bacillus thuringiensis
 Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus; Bacillus cereus group.
 1 (bases 1 to 3753)
 AUTHORS Thorne, L., Garduno, F., Thompson, T., Decker, D., Zouner, M., Wild, M., Walfeld, A. M. and Pollock, T. J.
 TITLE Structural similarity between the lepidoptera- and diptera-specific insecticidal endotoxin genes of Bacillus thuringiensis subsp. kurstaki and israelensis
 JOURNAL J. Bacteriol. 166 (3), 801-811 (1986)
 PUBMED 3011746
 COMMENT source text: B.thuringiensis subspecies israelensis DNA.
 FEATURES
 source location/Qualifiers
 1..3753
 /organism="Bacillus thuringiensis"
 /mol_type="genomic DNA"
 /db_xref="taxon:1428"
 941..2968
 /note="insecticidal endotoxin (put.); putative"
 /codon_start=1
 /transl_table=11
 /protein_id="AAA22614.1"
 /db_xref="GI:143229"
 /translation="NNPYONKKEVEIFNAPSGFSKSNYSRPLANKRNOPLKNTNYKDMANVCOONQYGNAGNAPASERTIVGASGIIYVGMGLAFAPVLAAGIISGTLIPFQSGDDPAAWMDLTNIGSRPIQEDIKKINIVTSIVTPIKQDLKYGQFPKWEPAITHANAAVAHDLFTTLEPIIDKQDLKNNASIRIPLPAVAGIATWHLULKHAATYNNITWLNQGINSTFNSNYYQGLKRLKIOEYDIYCIQTNAQITMIRITNATWNTNIRLEMTLTVLDLIAIPNYDEPKPIGVSKLIRBYNNVNSDFERTITTELENGLTRNPTLFWINQGFYTRNSRDLDPYDIFSETONMAFTHNDNRNIMCAVGNITSDPSKVPFPPRNKPIDKIVRARESDIIVEMLIPNSSEVERVSNSRTIENNYKRTDSYMIKQTKAREKGTHTSYITKDNIVISVABRRVARSWTHTSIDPNTTIDLDNITQIHAKLVKSSDSKIVKPGHTGDVLILKOSMPKRVKFLKNSKQIVRIRIYATNAPKTVFLGIDITISVELPSTTSRONPNADLYADFGVYTPRPVTPNKTFEBEDTILMTLYGTPNHSYNIYIDKIERIPITQSVLDYTEKONIEKTQKIVNDLFPV"
 ORIGIN
 Alignment Scores:
 Pired. No.: 7,08e-58 Length: 3753
 Score: 897.50 Matches: 247
 Percent Similarity: 48.3% Conservative: 126
 Best Local Similarity: 32.0% Mismatches: 270
 Query Match: 23.2% Indels: 129
 DB: 1 Gaps: 30
 US-10-783-417-2 (1-735) x BACMSOB (1-3753)
 Qy 1 MetAnglnAsnAsnAspAsnAsnGluTyrgluIleAspSerHisThrSerProTyR 20
 Db 941 ATGAATCCATATCAAAATTAAGAAATATGAAATTTCAATGCTCATCAATGTTT 1000
 Qy 21 PheProAsnArgAsnSerAsnAspSerArgTyProTyThrAsnAsnProAnglnPro 40
 Db 1001 -----AGCAAGCTTAATTAATTAATCTAGATTCATTAGCAAAATTAAGCAATCAACA 1054
 Qy 41 LeuGlnAsnThrAsnTyIleGluIleTyLeuAsnMetCysGlnGlyAsnThrGlnTyrgly 60
 Db 1055 CTGAAAACACACAAATTAACAAATGCGCTCAATGCTGTCAGATTAATCAAAATATGCGC 1114
 Qy 61 AspAsnPhgGluThrPheAlaSerAlaAspThrIleAlaIleValSerAlaGlyThrIle 80


```

Db      1115 AATATGCGGGGAATTTCCTAGTCTGAACTATGTTGGAGTACGAGGTATTTAT 1174
Qy      81 ValSerGlyThrLeuValaglyIleGlyLeuThrSerLeuSerGlyProIleGly 100
Db      1175 GTAGTAGAAGACATATGTTAGA-----GCTTTGGCTGCCCT----- 1210
Qy      101 IleIleGlyValIleIleSerPheGlyThrLeuIleThrValPheTyrProIleGly 120
Db      1211 GCTTAGGCTGCGAGTATATATCTTTGGGACTTTGGTCCGATCTTTTGG---CAAGGA 1267
Qy      121 GluGlnAspGlySerThrValTyrPheGlnPheIleGlySerMetGlyGluIlePheValAspThr 140
Db      1268 TCTGACCTCGCAATGTTTGGCAGAGATTTGTTAAACATCGA-----GGAAG 1315
Qy      141 ProLeuThrGlu-----SerIleGlyLeuLeuGlyLeuGlnThrLeuGluIleGlyPhe 157
Db      1316 CCTATACAGAAATATGATATTAACATATATATGATACCTTCTATGCTACCTTATA 1375
Qy      158 ArgGlnIleLeuGlnSerTyrAsnThrAlaLeuAspAspTyrArgGlySerLeuValArgLeu 177
Db      1376 AAAAATCAACTGATTAATATACAGAAATTTTTCGATTAATGGACGACGACACT----- 1429
Qy      178 GlnAlaProGlyLeuProProSerSerAlaLeuGlnGlnAlaIleLeuThrLeuValIle 197
Db      1430 -----ACACACGCTAATGCTAAAGCAGTACATGATCTTTTACT 1468
Qy      198 ArgPheGluAsnValHisAsn-----AspPheIleArgGluIleProGlyPheGln 214
Db      1469 ACCTTAGAACCTATATATGATATTAAGATTTAGATTTAAATAAATAAAGCTACGATGCA 1528
Qy      215 LeuGluThrTyrLeuThrLeuLeuProIleTyrAlaGlnAlaAlaAsnPheHisPhe 234
Db      1529 ATACCAACA-----CTCCCTGATATGACAAATAGCTACTTGGCACTTG 1573
Qy      235 AsnLeuLeuGlnGlnGlyValaGluLeuAlaAspGluTyrPheAlaAspIleHisProSer 254
Db      1574 AATTTATTAATAACATGCTGCTACTCTATTAACATATATG-----CTGCAAAATCA 1624
Qy      255 GlnIleGluProAsnAlaGlyThrSerAspAspTyrTyrLeu-----LeuLeuGlyGluAsn 273
Db      1635 GGTATTAATCCAAGTACTTCAATCATCTATATCTATACAGGCTATTTAAACGTATA 1684
Qy      274 IleProGlyTyrSerAsnTyrCysAlaAsnThrTyrArgThrGlyLeuValAsnLeu 293
Db      1685 ATACAGAAATATACGACTATTTGTATACAAACGATCAATGACGAGCTACATGATGATA 1744
Qy      294 AspGluProAsnMetLeuTyrSerIlePheAsnAspTyrArgArgTyrMetThrIleThr 313
Db      1745 ACTAATACTTAACGCAACATGAAATATGTAATATACCTTACCGTTTGAATATGCTTAACT 1804
Qy      314 ValLeuAspThrIleSerGlnPheSerLeuTyrAspIleValArgTyrArgAspSerIle 333
Db      1805 GGTATAGATCTTATTTGCTATTTTCCAAATTTATGACCCAGAAAAATAT----- 1852
Qy      334 GlyGlyIleGluValIleGlyIleLeuValAsnGluLeuThrArgGluIleTyrThrThrGlu 353
Db      1853 -----CCAATAGAGCTTAAATGTAATCTTATACGAGAACTTAT---ACGAAT 1897
Qy      354 IleAsnPheAspArgLeuProGlnLeuValGlnProAsnLeuAlaThrMetGluTyr 373
Db      1898 GTTAATCTCAGATCACTTT-----AGAACATTAACAGAACTAAGAAAT 1939
Qy      374 AsnLeuThrArgAlaSerPheLeuLeuPheSerPheLeuGlnIlePheIlePheTyrThr 393
Db      1940 GGATTAACCTAGA---AATCTCACTATTATTTACTTGATTAACCAAGGCGTTTACACA 1996
Qy      394 GluAsnThrAsn-----PheGlyAsnArgLeu 402
Db      1997 AGAATTTCTCGAGACATTTCTGATCTTATGATATTTTCTTTTACAGGTAAACAGATG 2056
Qy      403 ValGlyIleSerAsnArgAspAlaProThrTyrSerAsnThrIleThrGluThrLeuTyr 422

```

```

Db      2057 -----GCTTTACACTACTAATGATGTCGACAACTAATCTCG 2095
Qy      423 GlyGluArgThrGly-----SerProThrThrIleArgProPheGluSer 439
Db      2096 GGAAGCGTTCATGGAATAATATTTCTTCAGACACATCCAAAGTATTTCTTTTATAGA 2155
Qy      440 Tyr-----LysValSerIleValThrAspArgGlnSerProValSerPro 455
Db      2156 AACAAACCTATGATTAAGTCGAAATGTCAGACATAGACAGTACTCAGATATATA--- 2212
Qy      456 IleGlnProHisPheIleIleAsnGlnIleGluLeuTyrLeuAsnGlySerSerAsnAsn 475
Db      2213 -----TATATAATGATATTTTTCGAAATAGCGATGA 2245
Qy      476 ThrLeuSerTyrSerAlaGlyIleSerLeu---SerAsnTyrGlnAsnThrThrPhePhe 494
Db      2246 GTATTCGATATTTATTCATCCAAATTCACAAATATAGAAATATTTATTAAGAAACGATCTTAT 2305
Qy      495 GlnPheProArgGlyLeuAspCysAsnLeuValIleAspProGlyCysSerProAsnPhe 514
Db      2306 ATGATTCCAAAACAAACATGAAAAAT----- 2332
Qy      515 AsnAsnTyrSerHisIleLeuSerHisPheSerLeuPheThrTyrSerTyrValIleGly 534
Db      2333 GAAGAATATGATCACTATCTTATCGTATTAATAAATCTGATTAATATATTTTCACTAGCT 2392
Qy      535 LeuGlnLeuGlnIleLeuAspThrGlyValLeuGlyTyrPheHisSerSerValAspArg 554
Db      2393 AGAGAAAGAAAGAGGTT-----GCATTTAGTTGACACATACATAGTGTGATTTTC 2443
Qy      555 TyrAsnAlaIleSerAspGlyIleIleThrMetIleProAlaIleLeuGlyAsnAsnLeu 574
Db      2444 CAATAATCAATATGATTTGATATACATCACCCAAATCCACGCTCTAAAGCTTTGAAGGTA 2503
Qy      575 AspThrAsnSerLeuValIleGluGlyProGlyHisThrGlyGlyValAsnLeuValTyrLeu 594
Db      2504 AGTTCTGATGAAAAATTTGTAAGAGTCTCGTGCACACAGGTGGAGCTTGTAATTTCT 2563
Qy      595 GlnSerGlnGlyArgLeuGluIleThrCysGluThrProAsnSerThrGlnSerTyrPhe 614
Db      2564 AAAGATAGTATGAGATTTTAGAGTTAGATT---TTAAAAATGTTTCTCGCAATATTCGA 2620
Qy      615 IleArgLeuArgTyrAlaIleThrAsnGlyValaGlyAsnThrLeuProAsnIleSerLeuThr 634
Db      2621 GTACGATATTCGATATGCTACTAATGCT-----CCAAAGCAACAGTATATTC 2665
Qy      635 IleProGlyValIleGlyIleProProGlnArgLeuAsnAsnThrPheSerGlyThrAsn 654
Db      2666 TTAACCGGAATATGATATCTATAGTGTGAG---CTCCCTAGTACCATCTCCGCCAAAC 2722
Qy      655 TyrAsn-----AsnLeuGlnTyrGlyValAspPheGlyTyrPheGlnPheProSerThrVal 672
Db      2723 CCAATGCTACAGATTTAATACATATGCAAGATTTTGGATATGTAACTTTCCAAAGAACGTT 2782
Qy      673 -----ThrLeuProLeuAsnArgAsnIleProPheIle 683
Db      2783 CCAATTAACACTTTGAAGAGAGAGACCTTATTAATAGACC-----TTA 2827
Qy      684 PheAsnArgAlaAspValSerAsnSerIleLeuIleIleAspGlyIleGluPheIlePro 703
Db      2828 TATGTTACACCAACATCACTATATATATA---TATATTTGCAAAATAGAAATTTTTCGA 2884
Qy      704 IleThrSerSerMetHisGlnAsnArgGlyLeuArgGlyLeuGluValThrIleGlnThrLys 723
Db      2885 ATCACTCAATCGTATTTGATATATACAGAGAGCAAAATATATAGAAAAACACAGAAAAATA 2944
Qy      724 IleAsnThrPhePheThrAsnHisThrLysThrLeu 735
Db      2945 GTGAATGATTTATTTGTTAATTAACAAAGTCTT 2980

```

RESULT 11
BTPBTOXIS

BTPBTOXIS 127923 bp DNA linear BCT 16-APR-2005

DEFINITION Bacillus thuringiensis subsp. israelensis plasmid pBtoxis.
ACCESSION AL731825
VERSION AL731825.1 GI:21685410
KEYWORDS
ORGANISM Bacillus thuringiensis serovar israelensis
Bacillus thuringiensis serovar israelensis
Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus; Bacillus cereus group.
REFERENCE 1 (bases 1 to 127923)
AUTHORS Berry,C., O'Neil,S., Ben-Dov,E., Jones,A.F., Murphy,L., Quail,M.A., Holden,M.T., Harris,D., Zaritsky,A. and Parhill,U.
TITLE Complete sequence and organization of pBtoxis, the toxin-coding plasmid of Bacillus thuringiensis subsp. israelensis
JOURNAL Appl. Environ. Microbiol. 68 (10), 5082-5095 (2002)
PUBMED 12324359
REFERENCE 2 (bases 1 to 127923)
AUTHORS Parhill,U.
TITLE Direct Submission
JOURNAL Submitted (19-APR-2002) Submitted on behalf of the pBtoxis sequencing team, Sanger Centre, Wellcome Trust Genome Campus, Hinxton, Cambridge CB10 1SA E-mail: parhill@sanger.ac.uk
COMMENT Notes: Details of pBtoxis sequencing at the Sanger Centre are available on the World Wide Web.
FEATURES
source location/Qualifiers
1..127923
/organism="Bacillus thuringiensis serovar israelensis"
/mol_type="genomic DNA"
/db_xref="taxon:1430"
/note="plasmid pBtoxis"
1145..1149
/gene="pBt004"
1163..2089
/gene="pBt001"
1163..2089
/note="pBt001"
/note="Similar in part to Bacillus anthracis pXol-49 TR:Q9X319 (EMBL:AF065404) (227 aa) fasta scores: E(): 8.9e-44, 78.48% id in 158 aa"
/codon_start=1
/transl_table=11
/protein_id="CAD30064.1"
/db_xref="GI:21685411"
/db_xref="UniProt/TREMBL:O8KX5"
/translation="MSTHNFSCQNPVKHGVFPDGLNIIILENERKELPSALATY IMLAPEDNIGLPRFQIMDLAKSGIPIYTIYGFQVCLERKLVREIPVGNRYVE IDVALYNHTAETDMINISLSYRIPFYLIQTILSLVARNRGMILLETLYT FSRKIGMHYKHIEDFETTRMAFLKEKLNNAKROYIIVAPIKENAVDLKDK SSESRTYVRQVOQVITIEKPVIVSNVLENDKKELHPIAKCRKAVSRLLKMGQ ALRKDKENIMTAFROEIVDIIYIPTYKOKN"
2246..3106
/note="IS240"
/complement(2246..2262)
/rpt_type=INVERTED
2337..3044
/gene="pBt003"
2337..3044
/gene="pBt003"
2337..3044
/note="pBt003"
/note="Similar to Bacillus thuringiensis insertion element IS240-a protein TR:Q45766 (EMBL:M23740) (235 aa) fasta scores: E(): 2.2e-92, 99.57% id in 235 aa, and to Mycobacterium fortuitum, transposase tnp tnp6100 TR:Q49185 (EMBL:X53635) (254 aa) fasta scores: E(): 1.4e-37, 48.05% id in 231 aa"
/codon_start=1
/transl_table=11
/product="insertion element IS240 protein"
/protein_id="CAD30065.1"
/db_xref="GI:21685412"
/db_xref="GOA:O8KX4"
/db_xref="InterPro:IPR001584"
/db_xref="UniProt/TREMBL:O8KX4"
/translation="MEKENIFPKMKHYQADMILMTYRWYLRYNLSFRDLVEMMEEGLS

LSHTTMRVHGYGPELNERIRKHLKRTNDSNRVDETYIKIKGNMYLYRAVDSGNT LDPLYSKRDAKACFLKKALASPHYVTKPRVITYDGNAKAYVALRELKNSKIPYGM PLRVKYLNNMIEODHRIKRIILNMLGKSKQTAVKMIAIGIEAHMVKKGLKLRQ SAQONRCIHOIFGLTA"
2538..2933
/gene="pBt003"
/note="HMPFam hit to PF00665, Integrase core domain"
3090..3106
/rpt_type=INVERTED
3521..4381
/note="IS240"
/complement(3521..3535)
/rpt_type=INVERTED
3612..4319
/gene="pBt004"
3612..4319
/gene="pBt004"
/note="pBt004"
/note="Similar to Bacillus thuringiensis insertion element IS240-a protein TR:Q45766 (EMBL:M23740) (235 aa) fasta scores: E(): 3.5e-91, 99.14% id in 235 aa, and to Mycobacterium fortuitum, transposase tnpA or tnp6100 TR:Q49185 (EMBL:X53635) (254 aa) fasta scores: E(): 1.1e-37, 48.05% id in 231 aa"
/codon_start=1
/transl_table=11
/product="insertion element IS240 protein"
/protein_id="CAD30066.1"
/db_xref="GI:21685413"
/db_xref="GOA:O8KX5"
/db_xref="InterPro:IPR001584"
/db_xref="UniProt/TREMBL:O8KX5"
/translation="MEKENIFPKMKHYQADMILMTYRWYLRYNLSFRDLVEMMEEGLS LSHTTMRVHGYGPELNERIRKHLKRTNDSNRVDETYIKIKGNMYLYRAVDSGNT LDPLYSKRDAKACFLKKALASPHYVTKPRVITYDGNAKAYVALRELKNSKIPYGM PLRVKYLNNMIEODHRIKRIILNMLGKSKQTAVKMIAIGIEAHMVKKGLKLRQ SAQONRCIHOIFGLTA"
3813..4208
/gene="pBt004"
/note="HMPFam hit to PF00665, Integrase core domain"
4365..4381
/rpt_type=INVERTED
/complement(4634..5275)
/gene="pBt005"
/complement(4634..5275)
/gene="pBt005"
/note="Similar to N-termminus of Bacillus anthracis pXol-18 TR:Q9X2Y9 (EMBL:AF065404) (315 aa) fasta scores: E(): 4.2e-56, 84.15% id in 183 aa, and to Bacillus halodurans Bb2364 protein TR:Q9KAC5 (EMBL:AP001515) (378 aa) fasta scores: E(): 1.6e-18, 35.45% id in 189 aa, and weakly to Lactobacillus delbrueckii integrase/recombinase orf2 TR:Q48538 (EMBL:Z50864) (333 aa) fasta scores: E(): 6.3, 28.88% id in 90 aa, and to Bacillus thuringiensis resolvase tnpI SW:TNRI BACTU (P10020) (284 aa) fasta scores: E(): 8.5, 23.88% id in 180 aa"
/codon_start=1
/transl_table=11
/product="integrase/recombinase family protein"
/protein_id="CAD30067.1"
/db_xref="GI:21685414"
/db_xref="GOA:O8KX3"
/db_xref="InterPro:IPR004107"
/db_xref="UniProt/TREMBL:O8KX3"
/translation="MSKNINLEETLAASVAYLIEKRGQSTIKRYAYVIXDKPYKMLRA NKGLHIKSWSELSADYQTFSELEDRSKYQKTRHIVVLLKGLHMLCLVSPIDG INLSLIPDSINDNDPITEMERKILKOTVLSTKGLTEOAKRPLIMDNVCTIILVV NYGSLQRLVLSINSHINPARIWICYVKKIDTINDIIFVQGLIYAFHRTSF"
complement(4985..5251)
/gene="pBt005"
/note="HMPFam hit to PF02899, Phage integrase, N-terminal SAM-like domain"
complement(5283..5286)
complement(5364..5507)

```

CDS
/gene="pBt006"
/complement(5364..5507)
/gene="pBt006"
/notes="Similar to Bacillus anthracis pXol-17 TR:O9X2Y8
(EMBL:AF065404) (47 aa) fasta scores: E(): 2.1e-12, 68.08%
id in 47 aa"
/codon_start=1
/translation="putative integral membrane protein"
/protein_id="CAD30068.1"
/db_xref="GI:21685415"
/db_xref="UniProt/TREMBL:O8KXN2"
/translation="MAVLKDSNTYIMNTYNEIGFPGKIFWGLFVVPFMSIMATPIW
LCK"
/complement(5370..5435)
/gene="pBt006"
/notes="1 probable transmembrane helix predicted for pBt006
by TMHMM2.0"
/complement(5515..5519)
/gene="pBt007"
/complement(6451..8160)
/gene="pBt007"
/complement(6451..8160)
/gene="pBt007"
/notes="Similar to Bacillus anthracis pXol-16 TR:O9X2Y7
(EMBL:AF065404) (569 aa) fasta scores: E(): 0, 96.13% id
in 569 aa, and to Bacillus thuringiensis pXol orf10-like
protein TR:CA050562 (EMBL:AJ296638) (310 aa) fasta scores:
E(): 6.5e-122, 99.67% id in 310 aa"
/codon_start=1
/translation="conserved hypothetical protein"
/protein_id="CAD30069.1"
/db_xref="GI:21685416"
/db_xref="UniProt/TREMBL:O8KXN1"
/translation="MGKSFIIQNRSLNGKTNNAVVTLLKSMQFVAAPBKYYSKD
TYLIDLIQHTYERIRIADVEVERKAKSPSRASSTLAETGRGHKXGTTNADLYE
FLYLHKCFPTNGVIVNTITTHMMEDYKQKEFAYIQHSQFIALKJLNHTISTEN
GIDGRYTIKTLFPMETEKANPYVVISPVIFTKQFELSLAAKLFLIDIAQOHTET
TLKSLDKODERGNTHFGMYRPLAKYVPHQIRAVIELTALPTGNPLKICKMQ
KGVKHTKTYTLLYSIDPLCSKEAGEOHNDPTPATYARAKNFLEAVJOENQYI
ELADMKFTHVLYKTHCHROIRSVIRGLMDWRKSGPTKIVTLTKLILVQSYOI
LDTAKSGIYPLIADHIREKRSDBKQAVNFGILSYMSLHNTIRMRKNTALIKOK
FAPVTESTYRNTLYKQSEITLFRKATADQGNLHAYIALETEMEKIKVSKERTI
PSDVEWFIETADKLPOKRLVETLPOKFLNEFMRTERLVRAGVITTAPOOVLAHM
BIX"
/complement(8168..8173)
RBS
8947..8951
RBS
8959..10653
/gene="pBt009"
8959..10653
/gene="pBt009"
/notes="Similar to Bacillus anthracis pXol-14 TR:O9X2Y5
(EMBL:AF065404) (564 aa) fasta scores: E(): 1.6e-191,
Alignment Scores:
Pred. No.: 5,88e-56 Length: 127923
Score: 897.50 Matches: 247
Percent Similarity: 48.3% Conservative: 126
Best Local Similarity: 32.0% Mismatches: 270
Query Match: 23.2% Indels: 129
DB: 1 Gaps: 30
US-10-783-417-2 (1-735) x BTPBTOXIS (1-127923)
QY 1 MetAaGlnAAsuAaSpAaSuAaNgLuYrGluIlelleAaSpSerHieThSerProYr 20
Db 39061 ATGAATCCATATCAAAATATGAGATATGAAATATTCATCATGCTCCATCCATGCTTT 39120
QY 21 PheProAaAaGAsuAaSuAaSpSerAaGlyYrProYrThAAsuAaSuProAaGlnPro 40
Db 39121 -----ACGAAGCTAAATACATATCTAGATATCCATATGCAAAATATGACCAATACACCA 39174
QY 41 LeuGlnAaenThAaenTyrluGluIuTrleuAaenMetCyeGlnGlyAaenThrGlnTyrlu 60

```

```

Db 39175 CTGAAAAACAGAAATTAACAAAGATGGCTCAATGCTGCAGAAATACAAATATATGCG 39234
QY AaPaenPheGluThrPheAlaSerAlaAaPThrIleAlaAlaValSerAlaGlyThrIle 80
Db 39235 AATATATGCGGGGAATTTTCTGATTCGTAACATAATTTGGAGTTGACGATTAATTT 39294
QY 81 ValSerGlyThrLeuLeuAlaGlyIleGlyLeuThrSerIleSerGlyProIleGly 100
Db 39295 GTATAGAGAACTATATGTA-----GTTTGGCTGCCCCCT----- 39330
QY 101 IleIleGlyAlaIleIleIleSerPheGlyThrLeuIleThrValPheThrProAlaGly 120
Db 39331 GCTTAGTCGACAGATATATATATCTTTGGCACTTTGGCCATCTTTGG---CAAGGA 39387
QY 121 GluGlnAaPlyThrValTrpThrGlnPheIleLeuMetGlyGluIlePheValAaPThr 140
Db 39388 TCTGACCTTCGCAAAATGTTTGGCAGATTTGTTAAACATCGA-----GGAAGG 39435
QY 141 ProLeuThrGlu-----SerIleGlyGlnLeuYrLeuGlnThrLeuGluGlyPhe 157
Db 39436 CCTATACAAAGAAATAGATTAATAACATATTAATGACTTAACCTTATCGTACACCTATA 39495
QY 158 ArgGlnIleLeuGlnSerTyrrAaenThralaAaAaPThrPrArgYrLeuYrAaGlu 177
Db 39496 AAAAATCAACTGATTAATATCAAGAAATTTTCGATTAATGAGACCGACGACGT----- 39549
QY 178 GlnAlaProGlyLeuProProSerSerAlaLeuGlnAlaAlaIleuThrLeuYrIle 197
Db 39550 -----ACACACGCTAATGCTAAAGCAAGTACATGATCTCTTACT 39588
QY 198 ArgPheGluAaenValHieAaen-----AaPheIleArgGluIleProGlyPheGln 214
Db 39589 ACCTTAACAACTATATATATGATTAAGATTAAGTAAATAAATAATGCTAGTATCGA 39648
QY 215 LeuGluThrTyrluYrThleuLeuLeuProIleTyrlaAlaGlnAlaAlaAaenPheIleu 234
Db 39649 ATACCAACA-----CTCCCTGATATGACAAATAGTACTTGGACATGG 39693
QY 235 AaenLeuGlnGlnGlyValaGluLeuAlaAaPJuTrpAaenAlaAaPThrIleIleProSer 254
Db 39694 AATTTATTAATAACATGCTGCTACTTACTTATCAATATATATG-----CTGCAAAATCA 39744
QY 255 GlnIleGluProAaAlaGlyThrSerAaPThrTyrluYr-----LeuLeuYrGluAa 273
Db 39745 GGTATTAATCCAAAGTACTTCAATTCATCATCTATCTATACAGGCTATTTAAACGTAA 39804
QY 274 IleProGlyTyrrSerAaenTyrrCyalaAaenThrTyrrArgThrGlyLeuYrAaenLeu 293
Db 39805 ATACAGAAATATACTGACTATTTGTATACAAACGATGCAAGACTAATGATATGATA 39864
QY 294 AaPJuProAaenMetTyrrSerIlePheAaenAaPThrTyrrArgTyrrMetThrIleThr 313
Db 39865 ACTAATACTTAACGCAACATGGAATATGTAATATCTTACCGTTTGAATGACTTACT 39924
QY 314 ValLeuAaPThrIleSerGlnPheSerLeuTyrrAaPThrIleYrAaGlyTyrrAaPThrIle 333
Db 39925 GTGTATAGATCTTATTTGATTTTCCAAATATATGACCAAGAAATAT----- 39972
QY 334 GlyGlyIleGluValGlyIleYrAaenGluLeuThrArgGluIleTyrrThrGlu 353
Db 39973 -----CCAATGAGAGTTAAATCTGAATCTTACAGAGAGTTAT---ACGAAT 40017
QY 354 IleAaenPheAaPArgLeuProGlnLeuYrGluGlnProAaenLeuAlaThrMetGluTyrr 373
Db 40018 GTTATTTAGATACCTTT-----AAGACATATACAGAACTAAGAAAT 40059
QY 374 AaenLeuThrArgAlaSerPheYrLeuPheSerPheLeuGluGlnPheIlePheTyrrThr 393
Db 40060 GGATTTAACTAGA---AATCTTAATATTTTCTTGATTAACCAAGGGGTTTATACCA 40116
QY 394 GluAaenThrAaen-----PheGlyAaenAaGlu 402

```

Db 40117 AGAAATTCGAGACANTCTGANTCCTTAAGATATTTTCTTTACAGTAACGANG 40116
 QY 403 ValGlyIleSerAsnArgAlaProThrTySerAsnThrIleThrGluThrLeuTyR 422
 Db 40177 -----GCCTTACACATTAAGATGATGCAACAATATCTGG 40215
 QY 423 GlyGluArgThrGly-----SerProThrThySerThrIleArgProHepGluSer 439
 Db 40216 GAGGCGGTCACTGAAATATTTATTTTCCAAACACATCCAAATATTTCTTTTATAGA 40275
 QY 440 Tyr-----LysValSerIleValThrAspArgGluSerProProValSerPro 455
 Db 40276 AACCAACCTTATGATTAAAGTCCGAATTTGTACGACATRAGACAGTCTCAAGTATATA--- 40332
 QY 456 IleGlnProHisPheIleIleAsnGlnIleGluLeuTyLeuAsnGlySerSerAsnAsn 475
 Db 40333 -----TANGAATAGATATTTTTCGAATAGACGTGAA 40365
 QY 476 ThrLeuIleTySerAlaGlyGlySerLeu---SerAsnTyGlnAsnThrThrPhePhe 494
 Db 40366 GTATTTTGGAAATTCATCCAAATTCACAACTGAAATAATATTAAGAACTGATCTTAT 40425
 QY 495 GlnArgProArgLysLysAspCysAsnIleValIleAspProGlyCysSerProAsnPhe 514
 Db 40426 ATGATTCCAAAACCAACATGCAAAAT----- 40452
 QY 515 AsnAsnTySerSerHisIleLeuSerHisPheSerLeuPheThrTySerTyValIleGly 534
 Db 40453 AAAGAAATRATGGTACTACTCTATGGTATATAAACATGATATATATATTTTCAGTAGTT 40512
 QY 535 LeuGlnLeuGlnIleLeuAspThrGlyValLeuGlyTYrThrHisSerSerValAspArg 554
 Db 40513 AGAAGAAAGAAAGACAGTT-----GCATTTAGTTGGACACATCTAGTGTGATTC 40563
 QY 555 TyrTrsAsnIleIleSerAspLysIleIleIleThrMetIleProAlaIleLysGlyAsnAsnLeu 574
 Db 40564 CAAATATCAATAGATTTAGATTAACATCACCAATACCGCTTAAAGCTTTGAAGGTA 40623
 QY 575 AspThrAsnSerLysValIleGlnGlyProGlyVHISerThrGlyLysAsnLeuValTyLeu 594
 Db 40624 AGTTCTGATTCGAAATATGTGGAAGGTCCTGGTCAACACAGTGGAGACTTGGTAATCTT 40683
 QY 595 GlnSerGlnGlyArgLeuGlnIleThrCysGluThrProAsnSerThrGlnSerTyPhe 614
 Db 40684 AAAGATAGTAGATTTTGAAGATGATGATTT---TTAAAAATGTTCCTCGACATATATCAA 40740
 QY 615 IleArgLeuArgTyValaThrAsnGlyValGlyAsnThrLeuProAsnIleSerLeuThr 634
 Db 40741 GTAGTGAATGTGTATGTGTACTTAAAGCT-----CCAAAGACAAACAGATATTC 40785
 QY 635 IleProGlyValIleGlyIleProProGlnArgLeuAsnAsnThrPheSerGlyThrAsn 654
 Db 40786 TTAAACCGGAATAGTACTATTAAGTGGAG---CTCCCTTAGTACCACTTCGCGCCAAAC 40842
 QY 655 TyrAsn-----AsnLeuGlnIleTyGlyAspPheGlyTyPheGlnPheProSerThrVal 672
 Db 40843 CCAAAATGCTACAGATTTTAACATATGACGATTTTGATGATGTAAATTCACAAAGACAGTT 40902
 QY 673 -----ThleuProLeuAsnArgAsnIleProPheIle 683
 Db 40903 CCAAAATAAACATTTGAGAGAGACACATTTTATTAATGCC-----TAA 40947
 QY 684 PheAsnArgAlaSerValSerAsnSerIleLeuIleIleAspLysIleGluPheIlePro 703
 Db 40948 TATGGTACACCAAACTCATTCATATATAATA---TATATTAACAAATCCGAATTTATCC 41006
 QY 704 IleThrSerSerMetHisGlnAsnArgGluLysGlnIleLeuGlnIleThrIleGlnThrLys 723
 Db 41005 ATCACTCACTGTATTAAGATTATATACAGACAGACAAATATATGAAAAAACACAGAAATA 41064
 QY 724 IleAsnThrPhePheThrAsnHisThrIleThrLeu 735
 Db 41065 GTGAATGATTTATTTGTTAATTAACAAAGTCTT 41100

[illegible]

QY 81 ValSerGlyThrLeuLeuAlaGlyIleGlyGlyLeuThrSerIleSerGlyProIleGly 100
 DB 1175 GTATGAGAACTATGTAGGA-----GCTTTGCTGCCCT----- 1210
 QY 101 IleIleGlyAlaIleIleSerPheGlyThrLeuIleThrValPheTrpProAlaGly 120
 DB 1211 GCTTACGTCGACGATATATCTTTTGGACCTTTGTCGACCTTTTGG---CAAGCA 1267
 QY 121 GluGlnAspLysThrValTrpThrGlnPheIleLysMetGlyGluIlePheValAspThr 140
 DB 1268 TCTGACCCCTGCAAAATGTTTGGACAGATTGTTAAACATCGA-----GGAAGG 1315
 QY 141 ProLeuThrGlu-----SerIleLysGlnLeuLysLeuGlnThrLeuGlnGlyPhe 157
 DB 1316 CCTATACAGAAATAGATATAAAACATTAATTAATGTAACCACTTCTATGCAACCTATA 1375
 QY 158 ArgGlnIleLeuGlnSerLysThrAsnThrAlaLeuAspAspTrpArgLysLeuLysArgLeu 177
 DB 1376 AAAAATCAACTGATTAATATCAAGAAATTTTTCGATTAATGGAGCCAGCACGCT----- 1429
 QY 178 GlnAlaProGlyLeuProProSerSerAlaLeuGlnGlnAlaAlaLeuThrLeuLysIle 197
 DB 1430 -----ACACACGCTATGCTAAAGCAGTACAGATCTCTTACT 1468
 QY 198 ArgPheGluAsnValHisAsn-----AspPheIleArgGluIleProGlyPheGln 214
 DB 1469 ACCTTGAACTTATATAGATTAAGATTAGATGTTAAAAAATGATGCTAGCTATGCA 1528
 QY 215 LeuGlnThrTrpLysThrLeuLeuLeuProIleTrpAlaGlnAlaAlaAsnPheHisLeu 234
 DB 1529 ATACCAACA-----CTCCCTGCATATGACAAATAGCTATCTGGACATCTG 1573
 QY 235 AsnLeuLeuGlnGlnGlyAlaGluLeuAlaAspGluTrpAsnAlaAspIleHisProSer 254
 DB 1574 AATTATTATTAACACATGCTGCTACCTATTCATCATATATG-----CTGCAAAATCA 1624
 QY 255 GlnIleGluProAsnAlaGlyThrSerAspAspTrpTrpLys--LeuLeuLysGluAsn 273
 DB 1625 GGTATTAATCCAAAGTACTTCAATCATCTATCATATCAAGCGCTATTTAAAAAGCTAAA 1684
 QY 274 IleProLysTrpSerAsnTrpCysAlaAsnThrTrpArgThrGlyLeuLysAsnLeuArg 293
 DB 1685 ATACAGAAATATACGACTATGTATACAAACGTAACATGACAGCACTATGATATGAA 1744
 QY 294 AspGluProAsnMetLysTrpSerIlePheAsnAspTrpArgArgTrpMetThrIleThr 313
 DB 1745 ACTATATACTAACGCAACATGGAATATGATATATCTACCTTGAATATGACTTAACT 1804
 QY 314 ValLeuAspThrTrpIleSerGlnPheSerLeuTrpAspIleLysArgTrpArgAspSerIle 333
 DB 1805 GTGTATGATCTTATTTGCTATTTTCCAAATTATGACCCGAAAAATAT----- 1852
 QY 334 GlyGlyIleGluValLysGlyLysLeuAsnGluLeuThrArgGluIleTrpThrGlu 353
 DB 1853 -----CCAAATGAGAGTTAAATCTGAACCTTATCAGAGAAAGTTAT---ACGAAT 1897
 QY 354 IleAsnPheAspArgLeuProGlnLeuArgValGlnProAsnLeuAlaThrMetGluTrp 373
 DB 1898 GTTATATGATATCTT-----AGAACATTAACAGAACTATAAAAAT 1939
 QY 374 AsnLeuThrArgAlaSerPheLysLeuPheSerPheLeuGlnGlnPheIlePheTrpThr 393
 DB 1940 GGATTAACATAGA---AATCTACATTATTTACTTGATTAACCAAGGGGGTTTTTACACA 1996
 QY 394 GluAsnThrAsn-----PheGlyAsnArgLeu 402
 DB 1997 AGAAATTTCTCGAGACATCTTGATCTTATGATATTTTCTTTTACAGGTAAACAGATG 2056
 QY 403 ValGlyIleSerAsnArgAspAlaProThrTrpSerAsnThrIleThrGluThrLeuTrp 422
 DB 2057 GCCTTACACATATCTAATGATAT-----CGCACATATATCTGGACGAGCGTTCAAT 2107
 QY 423 GlyLysArgThrGlySerProThrThrLysThrIle-----ArgProPhe 437

DB 2108 GGAATATATTATCTCAAGACACATCCAAAGTATTTCTTTTATAGAAACAAACTATTT 2167
 QY 438 GluSerTrpLysValSerIleValThrAspArgGlnSerProProValSerProIleGln 457
 DB 2168 GAT-----AGGTCGAATTTGTCAACATAGAGATCTCGATATATAT----- 2212
 QY 458 ProHisPheIleIleAsnGlnIleGluLeuTrpLeuAsnGlySerSerAsnAsnThrLeu 477
 DB 2213 -----TATGAATGATATTTTTCGAAATAGCAGTAAGTATTT 2251
 QY 478 LysTrpSerAlaGlyGlySerLeu---SerLysTrpGlnAsnThrPhePheGlnPhe 496
 DB 2252 CGATATTCATCAATTCAACATAGAAAATATATATAAAGAACTGATTTCTTATATGATT 2311
 QY 497 ProArgLysLysAspCysAsnLeuValIleAspProGlyCysSerProAsnPheAsnAsn 516
 DB 2312 CCAAAACCAACATGGAATAT-----GAAAGA 2338
 QY 517 TrpSerHisIleLeuSerHisPheSerLeuPheThrTrpSerTrpValIleGlyLeuGln 536
 DB 2339 TATGTCATCTCTATCGTATATATAAACTGATATATATATATTTTCAGTATGAGAA 2398
 QY 537 LeuGlnIleLeuAspThrGlyValLeuGlyTrpThrHisSerSerValAspArgTrpAsn 556
 DB 2399 AGAAGAAAGATT-----GCATTTAGTTGGACACACTACTAGTTGATTTCCAAAT 2449
 QY 557 AlaIleSerAspLysIleIleThrMetIleProAlaIleLysGlyAsnAsnLeuAspThr 576
 DB 2450 ACAATAGTTTATGATTAACATCAACCAATCCACGCTCAAAAACCTTAAAGTAAGTTCT 2509
 QY 577 AsnSerLysValIleGluGlyProGlyValIleThrGlyGlyAsnLeuValTrpLeuGlnSer 596
 DB 2510 AATTCAAAAATTTGGAAGAGCCCTCGCTCACACAAGTAAAACTTGTTATTTCTTAAAGAT 2569
 QY 597 GlnGlyArgLeuGluIleThrCysGluThrProAsnSerThrGlnSerTrpPheIleArg 616
 DB 2570 AGTATGAATTTTAAAGTTAGATT-----TAAAAAATGTTTCTCAACATATCAAGTACGT 2626
 QY 617 LeuArgTrpAlaThrAsnGlyValaGlyAsnThrLeuProAsnIleSerLeuThrIlePro 636
 DB 2627 ATTGTTATGCTTACTTAATGCT-----CCAAAGACACAGTATCTTAAAC 2671
 QY 637 GlyValIleGlyIleProProGlnArgLeuAsnAsnThrPheSerGlyThrAsnTrpAsn 656
 DB 2672 GGAATATGATCTATTAAGTGTGAG---CTCCCTAGTACCACTTCCGCGCAAAACCCAAT 2728
 QY 657 -----AsnLeuGlnTrpGlyAspPheGlyTrpPheGlnPheProSerThrValThrLeu 674
 DB 2729 GCTACAGATTTTAAACATATAGCAGATTTTGGATATGTAACATTTTCAAGAACGTTTCA--- 2785
 QY 675 ProLeuAsnArgAsnIle-----ProPheIle-----PheAsnArg 686
 DB 2786 -----AATTAACCACTTTGAAGAGAGACACTTATTAATGACCTTTATATATGCTCA 2839
 QY 687 AlaAspValSerAsnSerIleLeuIleIleAspLysIleGluPheIleProIleThrSer 706
 DB 2840 CCAATATCTTCAATATATATAT---TATATGACAAAATATGATTTATTCATCAATCACTCA 2896
 QY 707 SerMetHisGlnAsnArgGluLysGlnLysLeuGluThrIleGlnThrLysIleAsnThr 726
 DB 2897 TCTGATATGATTTATACAGAGAAAGCAAAATATATGAAAAACACAGAAAATAGTGAATGAT 2956
 QY 727 PhePheThrAsnHisThrLysThrLeu 735
 DB 2957 TTATTTGTTATTAATAACAAAGTTCTT 2983
 RESULT 13
 AF285775 4959 bp DNA linear BCT 16-AUG-2000
 LOCUS Bacillus thuringiensis subsp. finitimus parasporal inclusion
 DEFINITION protein Cry (cry) gene, complete cds.
 ACCESSION AF285775

Db 936 ACAGTATAGCAGCCTATA-----ATACCTATCTTGGCCT 974
 Qy 119 ALaGluGluGlnAspLeuThrValTrpThrGlnPheIleuSerMetGlyValIlePheVal 138
 Db 975 TCT-----CAAGAAATATATCTTCCCGATTAATCTTACGATTAATTCAGAACCAACCTT 1028
 Qy 139 AspThrProLeuThrGlnSerIleuGluIleuValLeuGlnThrLeuGluGlyPheArg 158
 Db 1029 TATTCATTTCAGATCAAGAGGTTAGAGAGATGACATTACGCCATTGAAAGTTAAAA 1088
 Qy 159 GlnIleLeuGlnSerIleuThrValLeuAspAspThrParglyLeuValLeuGln 178
 Db 1089 GATAGTGAATAATATTGGAATAATCTTTCACATTGGATTAAT----- 1133
 Qy 179 AlaProGlyLeuProSerSerAlaLeuGlnGlnAlaIleuThrLeuValLeuArg 198
 Db 1134 -----AATCTTAATTCACAC-----AACACAACTACCTGACGAAAGA 1172
 Qy 199 PheGluAsnValHisAsnAspPheIleArgIleuProGlyPheGlnLeuGlnThrTrp 218
 Db 1173 TTCGAAGAGGTAAAGTGAATTTGTAGCGACGATGCCCTTCTTACGAAAGACTAT 1232
 Qy 219 LysThrLeuLeuLeuProIleuValIleGlnAlaIleAsnPheIleLeuAsnLeuGln 238
 Db 1233 GAACCAATACCTGTTATCACTTATGACAGCTGACAGTTACATTCTTCATTACGG 1292
 Qy 239 GlnGlyAlaGluLeuAlaAspGluTrpAsnAlaAspIleHisProSerGlnIleGluPro 258
 Db 1293 GATGGAATTAAGTGAATAATGAAATTAATTAAGTGAACAGGAGTACATG----- 1346
 Qy 259 AsnAlaGlyThrSerAspAspTrpTrpIleuLeuValLeuGlnIleProValSer 278
 Db 1347 -----CCGGGGATTTACTTTACAAAGAAATTTAAACAATATGCAATGAATAATATT 1397
 Qy 279 AsnTrpCysAlaAsnThrTrpArgThrGlyLeuValLeuAsnLeuArgAspGluProAsnMet 298
 Db 1398 GAGCATTTATTAAGTGTACATGAAGATTGTCTCTTTCTTAA-----AGCTAGAGTCA 1454
 Qy 299 LysTrpSerIlePheAsnAspTrpArgTrpMetThrIleThrValLeuAspThrIle 318
 Db 1455 AATGGCTAGGAATACATCAATATCGTACATTCTTAAACAGCTCTGATATGATGTCA 1514
 Qy 319 SerGlnPheSerLeuTrpAspIleuValArgTrpArgAspSerIleGlyValIleGluVal 338
 Db 1515 TCTCTTTTTCATCATATGACCCAGTTTATTAAGAAAGATTA--AGTGTGACAGAT 1571
 Qy 339 LysGlyIleuValLeuThrArgGluIleuTrpThrGlnIleAsnPheAspArg 358
 Db 1572 -----CTCAGAGAAATTAATTAATACAGATCCATTAATTAATATATGA 1613
 Qy 359 LeuProGlnLeuArgValGlnProAsnLeuAlaThrMetGlyTrpAsnLeuThrArgAla 378
 Db 1614 GGCAATTCACCTAGAGAGAGTGAAGCAATATACCTTGAAGCCAACTTTGTTACTCA 1673
 Qy 379 SerPheIleuLeuSerPheLeuGlnPheIlePheTrpThrGlnAsnThrAsnPhe 398
 Db 1674 CTATATACACTTAACCTTTATATATATTTTAAATTAATCATGAGGACACACCAAT--- 1730
 Qy 399 GlyAsnArgLeuValGlyIleSerAsnArgAspAlaProThrTrpSerAsnThrIleThr 418
 Db 1731 -----ACTATAGATTAATTAATCAACCG 1751
 Qy 419 GluThrLeuTrpGlyValArgThrGlySerProThrThrTrpThrIleArgProPheGlu 438
 Db 1752 GACAAATTTTTCAGACAGGTTCTTTGGAAAAACAAGACAGCTAATATATGAT----- 1802
 Qy 439 SerTrpValSerIleValThrAspArgGlnSerProProValSerProIleGlnPro 458
 Db 1803 -----AAAGTCCCTGATATCTTAACGATTAATCA----- 1832
 Qy 459 HisPheIleIleAsnGlnIleGluLeuValLeuAsn----- 470

Db 1833 -----ATTATTTATTAATTCAGACATATGATATATCAATTAATGTTATTAATGTAATG 1886
 Qy 471 -----GlySerSerAsnAsnThrIleuValTrpSerAlaGlySer 484
 Db 1887 TACTTGGTTTGGACGAGGAAAGATCAATTAACAAATTAATTAATGAGTATCTT 1946
 Qy 485 LeuSerAsnTrpGlnAsnThrThrPhePheGlnPheProArglyValAspCysAsnLeu 504
 Db 1947 ACAGAAATATAC-----ATTAATAATTTGACT--- 1973
 Qy 505 ValIleAspProGlyCysSerProAsnPheAsnThrSerHisIleLeuSerHis--- 523
 Db 1974 -----CATCGCTTACGAGATGTCATCTCATCATAT 2003
 Qy 524 -----PheSerLeuPheThrTrpSerTrpValIleGlyLeuGlnLeuAsp 541
 Db 2004 CTGATGAAAAAACAATGTTATCTTT----- 2033
 Qy 542 ThrGlyValLeuGlyTrpThrHisSerSerValAspArgTrpAsnAlaIleSerAspLys 561
 Db 2034 -----GCATGACATCTACATCAATTTCTTGAATAATGAATTAATAATGAT 2081
 Qy 562 IleIleThrMetIleProAlaIleuGlyValAsnAsnLeuAspThrAsnSerLysValIle 581
 Db 2082 ATTATACACAAATCCCTGCTGTAAGACCTATCAATTAAGTGTGTCATCAACAGTAT 2141
 Qy 582 GlnGlyProGlyHisThrGlyValAsnLeuValTrpLeuGlnSerGlnGlyArgLeuGlu 601
 Db 2142 AAAGTCCAGCTCATACCTAGCTGAGACCTTAATTAATTAATGAATGATTAATTAATGCT 2201
 Qy 602 IleThrCysGlu---ThrProAsnSerThrGlnSerTrpPheIleArgLeuArgTrpAla 620
 Db 2202 ATATCATATGTCACACTTATCTTAATGTTACAAAAAATATTTTGTAAGATTCGCTACGCT 2261
 Qy 621 ThrAsnGlyAlaGlyAsnThrLeuProAsnIleSerLeuThrIleProGlyValIleGly 640
 Db 2262 ACAATGTTCTTAATATACAGGCCATTAATTAATTAATTAATCAATTCGA-----GSA 2312
 Qy 641 IleProProGlnArg-----LeuAsnAsnThrPheSerGlyThrAsnTrpAsnAsnLeu 658
 Db 2313 ATGACACCAACAGGATGATGCTCATATATCTTTCTGATACAGCTTAATAGCAATTTA 2372
 Qy 659 GlnTrpGlyAsnPheGlyTrpPheGlnPheProSerThrValThrLeuProLeuAsnArg 678
 Db 2373 GAAATATCAAAATTTGGCTATTAAGAAATTTCTAAAGAGTACATTAATCCAAATCA 2432
 Qy 679 AsnIleProPheIlePheAsnArgAlaAspVal---SerAsnSerIleLeuIleLeuAsp 697
 Db 2433 AGTATATCTCTTACATTAATCGTTTCATCAAAATTCAAATTCATTAATCTCTCTGAT 2492
 Qy 698 LysIleGluPheIleProIleThrSerSerMetHisGlnAsnArgGluValGlnLysLeu 717
 Db 2493 AGAATTTGAATTTCTACCAATTAATCTCCCTTATACAAAGATTAAGAAACAACAAATTTA 2552
 Qy 718 GluThrIleGlnThrIleValIleAsnThrPhePheThrAsnHisIleThrValLeu 735
 Db 2553 GAAAAAATTCAGAAAAACGTAAACAGTTATTTTCAATTAATAATTAAGGGTTG 2606
 RESULT 15
 AF132928 4896 bp DNA linear BCT 14-JAN-2000
 LOCUS Bacillus thuringiensis subsp. finitimus Cry28a1 delta-endotoxin
 DEFINITION gene, complete cds.
 ACCESSION AF132928
 VERSION AF132928.1 GI:4574729
 KEYWORDS
 SOURCE
 ORGANISM
 . Bacillus thuringiensis serovar finitimus
 Bacillus thuringiensis serovar finitimus
 Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus; Bacillus
 cereus group.
 1 (bases 1 to 4896)
 REFERENCE
 Wojciechowska,J.A., Lewitin,E., Revina,L.P., Zalunin,I.A. and
 Chestukhina,G.G.

TITLE Two novel delta-endotoxin gene families cry26 and cry28 from
JOURNAL *Bacillus thuringiensis* ssp. *finitimus*
PUBMED 10403372
REFERENCE 2 (bases 1 to 4896)
AUTHORS Wojciechowski, J.A., Lewitin, E.I. and Cheshchukina, G.G.
TITLE Direct Submission
JOURNAL Submitted (04-MAR-1999) Laboratory of Protein Chemistry, Institute
of Microbial Genetics, Dorozhny proezd, Moscow 115545, Russia
FEATURES
Location/Qualifiers
1..4896
/organism="Bacillus thuringiensis serovar finitimus"
/mol_type="genomic DNA"
/strain="VKPM B-1161"
/sub_species="finitimus"
/db_xref="taxon:29337"
1129..4458
/codon_start=1
/transl_table=11
/product="Cry28Aa1 delta-endotoxin"
/protein_id="AAD24189.1"
/db_xref="GI:4574730"
/translation="MAOTYYKIGVOSTEVSSEIIFNPEDVSDPTVAVAGIYVGC
ILTAFAFVNPGVILISRGTLAPVTPPPEEDPKIKNGPKHSGEDILNQTSTAVKE
IALAHNGFKOVLTTERAFNDWKNSPANTARLVSOEFNAHFVSNMPOLOPTY
DTLLSCYTEAANLNLHQGVQPADMDQPHSPMLKSGTYBELVIEKYIN
YCTKYHGLNHLKESKIKTWDVATYREMTLLVLDVATFPYDIRPFGVELT
TREYVTSLDHLTRPGFLFTMLSDIELYTESVABGVLSGTYEYVQNPFTKNIY
GNTRLSKOLITLIPGFMTLSINRPGTLINQKYSLSLOKIVFTPKNDYKOKY
FNVNQNPQETTYPNPDYGSNSQKPKHNSHPILTHKEPAFTHSTALQMTN
SVNSQNLISBSVSTQPLVKAIEVTVNSVINGPFTGDLIELDKCSIKKASLKK
YALSIFYANNAIVASIDVDSGAVLLQPTFSRGNNFTIODLVNDFCKKSLVD
IELPSESEIHILKQEDYERGVILLIDLEKFKPIDENYTEMNLEKAKKAVNLFN
ATNALMKQVTDHIDQVANLVEISDILYAKKIKFTFCISFKQLSQARMLPDL
NNIAENSTANTGVTTIEGDPVYKRAIOISAADENFPYLYOKIDESILKYTRY
OLRGVESGSDLELDVYRGATDYMANPQGLIELISYAPINPCEIETRDITTCGAL
DRKQSNVNSADVDVPQVNGDPHAFPHIDTGTNNRNLGIWIKIATPQGYAT
FGNLELIELGPLSGEALAOVQKEQKWKNTQKREBAKLYAAKQVINDYFPLD
TKLRFPIESNILSADKLVIKIDYSEVLVIPGLNATDLPMELENRIOMNIDLYDAR
NTVTGSEFRNGLANMASNTSEVRIOAHPCMYISGNMAQVAVNLEPDKGYLART
AKKGIQNGYVLTILDCANHIDTLTPSSCDSEFTTSSNELAIVYKTEILFPDITQIRI
EIGETRSIFYESVDLIMED"

ORIGIN

Alignment Scores:

Pred. No.:	4,62e-54	Length:	4896
Score:	849.00	Matches:	229
Percent Similarity:	48.84	Conservative:	106
Best Local Similarity:	33.34	Mismatches:	248
Query Match:	21.94	Indels:	104
DB:	1	Gaps:	25

US-10-783-417-2 (1-735) x AF132928 (1-4896)

QY 68 SerAlaAPThrIleAlaIaValSerAlaGlyThrIleValSerGlyThrLeuLeuAla 87
Db 1210 AGCAGTGTATACGTCGTGTAGTAAAGCCAGGAGATTGATGTCGGTACTATACAGCA 1269
QY 88 GlyIleGlyGlyLeuThrSerIleSerGlyProIleGlyIleIleGlyAlaIleIleIle 107
Db 1270 -----GCCCTTGACATTTGTTAATCA-----GGTGTGCTACTTATA 1308
QY 108 SerPheGlyThrLeuIleThrValPheTrpProAlaGlyGluGluAsp---LysThrVal 126
Db 1309 TCATTGGAACTTGCGTCGCCGTTCTTGCGCTGATCCAGAGAGATCCAAAATAATT 1368
QY 127 TrpThrGlnPheIleLysMetGlyGluIlePheValAspThrProLeuThrGluSerIle 146
Db 1369 TGTGCACATTTATGAAACACGAGAGACCTTTAATAACAAACATTTCTACAGCTGTA 1428
QY 147 LysGlnLeuLysLeuGlnThrLeuGluGlyPheArgGlnIleLeuGlnSerTyraThr 166
Db 1429 AAAAGAAATGACATTGCTCATCTAATAATGTTTAAAGATGTAATTAACGTACTATGAAAGA 1488

QY 167 AlaLeuAPAPThrPArgGlyLeuLysArgLeuGlnAlaProGlyLeuProSerSer 186
Db 1489 GCATTATATGATGGAGAGAC-----ATCCAAAGTCGA 1521
QY 187 AlaLeuGlnGlnAlaIleLeuThrLeuLysIleArgPheGluAsnValHisAsnAspPhe 206
Db 1522 -----AATTAAGTCGAGATGGTGTACACAGAGATTTGAAACGCTCATTTCAATTT 1572
QY 127 IleArgGluIleProGlyPheGluLeuGluThrTyrrTyrrThrLeuLeuProIleTyrr 226
Db 1573 GTAAGCAATATGCCAACATCCAACTCCACGATGACACATTTATTTAAGTTGCTAT 1632
QY 227 AlaGlnAlaAlaAsnPheIleLeuAsnLeuLeuGlnGlnGlyAlaGluLeuAlaAspGlu 246
Db 1633 ACAGAGCTGCAAAATTCATTTGATTTATTTATTCATCAAGGATGATCAATTCGCGGATCA 1692
QY 247 TrpAsnAlaAspIleHisProSerGlnIleGluProAsnAlaGlyThrSerAspAspTyrr 266
Db 1693 TGGATGACAGATCAACACATTCACCAATGTTGAAGTCATCAGGTACT-----TAT 1743
QY 267 TyrrLysLeuLeuLysGluAsnIleProLysTyrrSerAsnTyrrCyAlaAsnThrTyrrArg 286
Db 1744 TATGACAGCTATTTGCTATATTTGAAAGATATTTATTTATTTGACCAAGACATACAT 1803
QY 287 ThrGlyLeuLysAsnLeuArgAspGluProAsnMetLysTrpSerIlePheAsnAspTyrr 306
Db 1804 AAAGATTTGATTCACCTTAAAGAAATCAGAAAATTCACATGAGATGCTTATATACATAT 1863
QY 307 ArgArgTyrrMetThrIleThrValLeuAspThrIleSerGlnPheSerLeuTyrrAspIle 326
Db 1864 CGTGAGAAATGACCTTAAATGATTTGATGATCTTCGCAACCTTTCTTTTATGATATTA 1923
QY 327 LysArgTyrrArgAspSerIleGlyGlyIleGluValLysGlyIleLysAsnGluLeuThr 346
Db 1924 CGTCGTTTT-----CCAAAGAGATGAACTGAATTTACGA 1959
QY 347 ArgGluIleTyrrThrGluIleAsnPheAspArgLeuProGlnLeuArgValGlnPro 366
Db 1960 AGAGAGCTTATACAGATTGATGATTCATTTAACAGACACACA----- 2001
QY 367 AsnLeuAlaThrMetGluTyrrAsnLeuThrArgAlaSerPheLysLeuPheSerPheLeu 386
Db 2002 -----GGGCTATTTACTTGCGCTG 2019
QY 387 GluGlnPheIlePheTyrrThrGluAsnThrAsnPheGlyAsnArgLeuValGlyIleSer 406
Db 2020 TCAATATTGATGATTATACACGAGAGTGTGCAGAGCGATTTATTTACAGATTT--- 2076
QY 407 AsnArgAspAlaProThrTyrrSerAsnThr-----IleThrGluThrLeuTyrrGly 423
Db 2077 ---CGAGGCTCAATATATTATACGTAATCAATTTTACAGTAAATAATTTATGCT 2133
QY 424 GluArgThrGly---SerProThrThrLysThrIleArgProPheGluSerTyrrLysVal 442
Db 2134 AATCAATATGATTAAAGTACAGCTCATTCATTTATACAGGCGAA---TTTATGACT 2190
QY 443 SerIleValThrAspArgGlnSerProProValSerProIleGlnProHisPheIleIle 462
Db 2191 CACTTACGATTAACCGTCTTTTCAAACAAATAGCTGATTAATTAAGTTATACAGTTTA 2250
QY 463 AsnGlnIleGluLeuTyrrLeuAsnGlySerSerAsnAsnThrLeuLysTyrrSerAlaGly 482
Db 2251 ATTCAAAAATCGATTCACAACTTTTAAACGATAT-----GATATACAAAATAAT 2304
QY 483 GlySerLeuSerAsnTyrr-----GlnAsnThrPhePheGlnPheProArgLys 499
Db 2305 TTTATATGATTAATCAAAATGAACCTTCAAGAAACTCA----- 2343
QY 500 LysAspCysAsnLeuValIleAspProGlyCysSerProAsnPheAsnAsnTyrrSerHis 519
Db 2344 -----AACTATCTTAATGATTAATGTTGCTTCA---AACAGCCAAAATAATCAACAT 2391

QY 520 IleuSerHisPheSerIeuPheThrTyrSerTyrValIleGlyLeuGlnLeuGlnIle 539
DB 2392 AATTATCTCATTTTCATTATTCATCCAC-----AAGTTAGAGTTT 2433
QY 540 LeuAsp-----ThrGlyValLeuGlyTyrThrHisSerSerValAspArgTyr 555
DB 2434 GGTGAGTATTTTCACCTCTATTTGTCATTAGGTGACACATATGTGTAAACTCCCA 2493
QY 556 AsnAlaIleSerAspLysIleIleThrMetIleProAlaIleLeuGlyAsnAsnLeuAsp 575
DB 2494 AATTAAATATCAGAAAGTGTAGTACACAAATCCCATGTGTAAAGCTTACGAAGTT--- 2550
QY 576 ThrAsnSerLysValIleGlnGlyProGlyHisThrGlyGlyAsnLeuValTyrLeuGln 595
DB 2551 ACTAACAAATTCAGTTATAGAGCACAGGTTTACAGTGAGATTATATAGAACTTCGT 2610
QY 596 SerGlnGlyArgLeuGlnIleThrCysGlnIleThrProAsnSerThrGlnSerTyrPheIle 615
DB 2611 GAT-----AAATGTTCTATTAATGTAAAGCT--AGTCTTTAAAAAAATACGCTATA 2661
QY 616 ArgLeuArgTyrAlaThrAsnGlyAlaGlyAsnThrLeuProAsnIleSerLeuThrIle 635
DB 2662 AGTCTATTTTATGCTGCAATACGCAATAGCTGTATCAATAGACGTAGTGATTCGGA 2721
QY 636 ProGlyValIleGlyIleProProGlnArgLeuAsnAsnThrPheSerGlyThrAsnTyr 655
DB 2722 GCAGGAGTTCTA-----TTGCAACCTACCTTTTCTAGAAAAAGGGAAC 2763
QY 656 AsnAsn-----LeuGlnIleTyrGlyAspPheGlyTyrPheGlnPheProSer 670
DB 2764 AATTAATTTTACATTCAGACCTTAACATTAAGATTTCATATCATACACTTTTACTT 2823
QY 671 ThrValThrLeuProLeuAsnArgAsnIleProPheIlePheAsnArgAlaAsp----- 688
DB 2824 GATATTGAATTACCGAAGTAGAAGAAATTCATATCCATTGAGCGAGAGAGATGATTAT 2883
QY 689 ValSerAsnSerIleLeuIleIleAspLysIleGluPheIleProIleThrSerSerMet 708
DB 2884 GAGGAGGAGGTGATTTTATTTAGTATTAATAGAGTTCAACCTTATAGTGAATAATTAT 2943
QY 709 HisGlnAsnArgGlnLysGlnLysLeuGlnIleThrIleGlnThrLysIleAsnThrPhePhe 728
DB 2944 -----ACTAATGAATGAATTAGAGAAAGCAAGAAAGCAAGTGAATGATATATT 2994
QY 729 ThrAsnHisThrLysThrLeu 735
DB 2995 ATTAACGCAACAACGCTTTG 3015

Search completed: January 20, 2006, 20:29:24
Job time : 9852 secs

THIS PAGE BLANK (USPTO)